2019

Who is Better and Who is Best? What Differentiates Stars from the Rest

Montana R. Woolley
Wright State University

WHO IS BETTER AND WHO IS BEST? WHAT DIFFERENTIATES STARS FROM THE REST

A thesis submitted in partial fulfillment of the
requirements for the degree of
Master of Science

By

MONTANA R WOOLLEY
B.S., Wright State University, 2016

2019
Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

NOVEMBER 19, 2018

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Montana R Woolley ENTITLED Who is Better and Who is Best? What Differentiates Stars from the Rest BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science.

Corey Miller, Ph.D.
Thesis Director

Scott N. J. Wattamaniuk, Ph.D.
Graduate Program Director

Debra Steele-Johnson, Ph.D.
Chair, Department of Psychology

Committee on
Final Examination

Gary N. Burns, Ph.D.

Nathan Bowling, Ph.D.

Barry Milligan, Ph.D.
Interim Dean of the Graduate School

ABSTRACT

Woolley, Montana R. M.S. Department of Psychology, Wright State University, 2019. Who is Better and Who is Best? What Differentiates Stars from the Rest.

Star employees have significant influences on the successes or failures of organizations. Current research on stars has not addressed who a star is or how stars are different from other good employees. In this study I tested the efficacy of a proposed definition of star employees and verified the accuracy of other previously established characteristics and behaviors associated with stars. In addition, I qualitatively explored managers' perceptions of star employees. The study consisted of two separate samples: managers identified on MTurk (n = 40) and high-level executives from various industries (n = 46). Participants provided a series of open responses and ratings of both a star employee and an above average non-star employee. Results show strong support for previous findings within the literature and offer additional information regarding the existence and importance of specific characteristics and behaviors which differentiate star employees from other high-performers. Overall, performance was found to be the most important differentiator of stars. Initiative, leadership, prosocial behaviors, and social skills were all found to be important star attributes.

TABLE OF CONTENTS

Page

LIST OF TABLES

I. INTRODUCTION

What do Bill Gates, Tom Brady, Elon Musk, Oprah Winfrey, Simon Biles, and Timothy Judge all have in common? They could all be considered a star. Stars are unimaginably productive, true rock stars, and incredibly successful. Stars can be found in every company and in every field. It's no secret that stars are valuable. Identifying star (or superstar) employees should be a critical element in any talent management , because star employees have a significant influence on the success or failure of an organization (Aguinis & O'Boyle, 2014). Talent is a critical source of competitive advantage and is often considered the primary cause of organizational success (Silzer & Church, 2010). Stars produce more than other employees, increase the productivity of those around them, and impact the performance of the organization (Aguinis & Bradley, 2015). Star employees are highly visible in their organization because they "generate exorbitant output levels and demonstrate superior performance in relation to other employees" (Oldroyd, 2012).

Current star performer identification methods rely on a manager's opinion to select which employees are stars. Some researchers have relied on both manager and peer ratings to overlap (e.g., Kelley & Caplan, 1993). Another method of identification is by a measurable output, such as sales dollars or number of publications. However, there is considerable debate about what aspect of performance a measure of output is really capturing. For example, this method may be capturing underlying productivity, social

capital, or even time spent traveling. Identifying star employees is necessary to fast-track and retain these employees as well as target star applicants during the selection process. To find a way to identify star employees before they become highly visible, researchers must discover what makes a star different from a good employee, or an-above average employee whom is not considered a star. The purpose of my study is to identify the characteristics, traits, and behaviors that differentiate a star from a good employee.

Star Employees

Existing research refers to star employees in many ways. The most common of these labels are star employees, star performers, top performers, scale tippers, difference performers, hyper performers, and game changers (Aguinis & Bradley, 2015). The definition of a star performer is also not agreed upon. Some researchers have defined stars by their performance (Aguinis & O'Boyle, 2014) while others identify aspects such as visibility (Groysberg, Lee, & Nanda, 2008). Below I discuss what it means to be a star employee.

Defining Stars

Many researchers define stars by their performance. Aguinis and O'Boyle (2014) define star employees by their high location on the production distribution. They claim that being a star is a relative position and can only be identified by comparing their productivity in relation to others' productivity (Groysberg, Lee, and Nanda, 2008; Rosen, 1991). This would mean that a star in one position and organization may not be a star in another. According to Aguinis and O'Boyle (2014) productivity is just a result-based measure of job performance. However, the level of performance required to become a

star is unclear. Most researchers do not define a specific level of performance, just that it should be disproportionate (e.g., Aguinis & O'Boyle, 2014; Groysberg, Lee, and Nanda, 2008; Nyberg, 2010). Beck, Beatty, and Sackett (2013) use one standard deviation difference, while Gagne (2000) uses top 10% (or about 1.28 standard deviations above the mean in a normal distribution). In addition to having high performance, researchers maintain that a star must be identified over time, meaning that one needs to see the higher performance over a period of time (Aguinis and O'Boyle, 2014). Aguinis and Bradley (2015) argue that stars not only produce more than others, but also increase the productivity of those around them and the performance of the organization.

Other researchers define star employees by their internal and external visibility. Groysberg, Lee, and Nanda (2008) claim that in addition to being disproportionately more productive, stars are also more visible in their labor market. Stars have an impact on their organization, so they receive more attention within the organization; therefore, they are internally visible. Stars also will have more attention from competitors, clients, or the media, making their high performance more public and observable; therefore, they are externally visible. When a high performer has internal and external visibility, they are considered a star (Oldroyd & Morris, 2012).

Though there is a lack of empirical evidence describing stars, there are an abundance of conjectural theoretic descriptions. Stars have been described as people who "get the *right* things done" and people who are "highly efficient" (Tartakosky, 2011; Daum, 2015). Stars are clear with priorities (Tartakosky, 2011; Steinert, 2013; Staffing, 2015) and know how to say "no" (Tartakosky, 2011; Steinert, 2013). They are problem

solvers (Tartakosky, 2011; Harnish, 2011; Staffing, 2015; Daum, 2015) and have the motivation and drive to always continue learning (Tartakosky, 2011; Maher; Staffing, 2015). Stars maintain a "can-do" positivity (Tartakosky, 2011; Steinert, 2013; Maher; Staffing, 2015) and are described as charismatic (Tartakosky, 2011; Harnish, 2011; Daum, 2015). Stars take initiative and are engaged in day-to-day activities (Steinert, 2013; Harnish, 2011; Maher; Daum, 2015). Stars have effective oral and written communication skills (Harish, 2011). Additionally, Tartakosky (2011) describes stars as intelligent, ambitious, detail-oriented, and highly disciplined. Tartakosky also maintains that stars stay on track, delegate well, plan day-to-day activities, and surround themselves with the *right* resources. These subjective descriptions are a potential future source for establishing traits and behaviors that may distinguish stars from other employees.

Call, Nyberg, and Thatcher (2015) attempted to develop a universal integrative definition of star employees, sourced from various research disciplines (e.g., economics, sociology, management, etc.). Call et al. (2015) define star employees as "those with disproportionately high and prolonged (a) performance, (b) visibility, and (c) relevant social capital." In this definition, *disproportionately high* means high relative to the star's peers and *prolonged* means sustained and not just a one-time success. Performance should encompass all aspects of the construct, including results-based performance and behavior-based performance. Visibility is the extent to which someone's reputation and job performance are observable (Merton, 1968a). As previously mentioned, a star can have both internal and external visibility. Visibility has many effects on the star. They likely would have more attention within the company and may be treated differently.

Additionally, it will likely affect the stars' mobility. Finally, stars have relevant social capital. This refers to a star's social network. Stars capitalize on valuable relationships—gaining resources and information from the relationships they maintain. Call, Nyberg, and Thatcher (2015) specify that the social capital needs to be relevant, meaning it needs to have strategic value. The researchers give three main reasons why social capital is so important to maintain stardom. First, great achievements and innovations are usually the result of many people collaborating. Second, it creates opportunities. Those with the appropriate network are more likely to be chosen for various opportunities. Finally, social capital is vital for gaining access to resources.

Across the star literature, one or more of these aspects are used to define star employees in the star employee literature: performance (including output and productivity), visibility, and relevant social capital. Productivity can easily be measured in jobs with a quantifiable output, such as sales. However, the majority of jobs do not have a quantifiable output which makes results-based measures difficult to utilize. Visibility, or how much an employee stands out within their labor market, does not have any established measurement techniques. The same goes for relevant social capital. Finally, attempts to measure job performance are abundant. Because job performance is applicable to all jobs and measurable, I will focus on job performance in this study. However, internal visibility and relevant social capital will be measured as well, using measures created for this study. If Call, Nyberg, and Thatcher's (2015) definition of star employees is accurate, then I would expect stars to be rated significantly higher on performance measures, the visibility measure, and the social capital measure.

Stars' Job Performance

Historically, job performance has been difficult to measure. Performance measures were a "criteria of convenience," meaning that researchers would use whatever measure was easiest to obtain (Campbell et al., 1993). Researchers eventually saw this criterion of convenience as a problem and began creating models to conceptualize the performance theories. For the past three decades work on defining and measuring performance has grown exponentially.

According to Campbell et al. (1993), job performance can be defined as any behavior that employees engage in while at work that contributes to organizational goals. Researchers maintain that performance is synonymous with behavior. Performance is not a result, but rather the action that causes the result. It is important to note that not all behaviors are observable, specifically cognitive processes which can only be seen by the result itself. Because of this problem, Campbell et al. (1993) argue that performance consists of goal-relevant actions that are under control of the individual. These researchers also specify the difference between performance, productivity, and effectiveness. Effectiveness is defined as the evaluation of the results of performance. Therefore, performance and effectiveness are related. The researchers argue that this "bottom-line," although very important, should not be used to measure performance because it is not under direct control of the incumbent. Productivity is defined as the ratio of effectiveness to the cost of achieving that level of effectiveness, therefore, it is the relationship between inputs and outputs. This is also related to job performance by

measuring how well a person or group is functioning. Cambell et al.'s (1993) definition of job performance allows behaviors that are not directly associated with task performance to be included as part of the job performance construct.

Campbell, McCloy, Oppler, and Sager (1993) made a distinction to separate task (in-role) performance and contextual (extra-role) performance. Task performance refers to a person's performance on tasks that are part of her technical job, whereas contextual performance refers to a person's performance in areas that do not directly relate to technical job tasks (Borman & Motowidlo, 1993). For example, when a barista at a coffee house brews a new pot of coffee this is task performance. The barista's ability to make the coffee is what would be assessed. The barista would be displaying contextual performance when she sees a new employee struggling to use the cappuccino machine and then goes over to assist the new employee. This helpful behavior did not relate to the barista's technical job tasks, but was still a behavior that contributed to the organizational goals of the coffee house.

Borman and Motowidlo (1993) explain that there are four aspects that make contextual performance different that task performance. First, contextual activities do not contribute to the technical job aspects but instead support the organizational, social, and psychological work environment. Second, contextual activities are common to many or all jobs, whereas task activities vary from job to job. Third, task performance varies with knowledge, skills, and abilities, whereas contextual performance varies with volitional and predispositional variables. Fourth, contextual activities are not role prescribed, and are generally not included in an incumbents list of formal responsibilities.

Campbell et al. (1993) presented a multi-dimensional model of job performance that was composed of eight factors that each separately influence job performance: (1) job-specific-task proficiency, (2) nonjob-specific-task proficiency, (3) written and oral communication, (4) demonstrating effort, (5) maintaining personal discipline, (6) facilitating peer and team performance, (7) supervision/leadership, and (8) management/administration. Campbell et al. (1993) intended this model to represent the performance domain of all jobs, though the researchers suggested that some dimensions may not be relevant to every job or the degree of salience may vary. Campbell et al. (1993) argued that only three of these dimensions are relevant to every job: core task proficiency, demonstrating effort, and maintaining personal discipline. Alternatively, Viswesvaran (1993) examined performance results across ten dimensions and found positive correlations across the dimensions. This result suggested that there is a general factor accounting for significant variance across virtually all dimensions of performance. Though the debate between a unidimensional model of job performance and a multidimensional model of job performance is still ongoing today, there has been empirical support for a multidimensional model (Tubre, Arthur, & Bennett, 2006). Once job performance has been defined it is then necessary to measure it. Murphy (1989) identified eight different ways that performance can be measured. These eight ways include paper and pencil tests, job skills tests, on-site hands-on testing, off-site hands-on testing, high fidelity simulations, symbolic simulations, task ratings, and global ratings. The most common of these eight methods are task and global ratings, typically completed by supervisors of the incumbent being assessed. Task and global ratings can

be useful for different purposes. For example, global ratings are usually best for making administrative decisions or succession planning. Task ratings are best for employee development and providing specific performance feedback. It is important to consider that every measurement method contains many potential sources of error. Researchers have taken steps to reduce this error, such as utilizing multiple raters, anchored rating scales, and objective measures. However, much error still exists in performance measurements.

Based on this research, I will use a global rating of performance completed by managers in this study. An overall unidimensional measure of performance will be used as well as several multidimensional areas of performance.

Distinguishing Stars from Others

Because high levels of job performance are important to organizations, there has been considerable research on this topic. In addition to research on stars, organizational researchers have identified other types of high performing employees. In this section stars are differentiated from other distinctions that have been made. While star's may certainly overlap with some of these other employee types, there is a difference in definition and identification within the literature.

High-potential employees.

There are a few areas of research that seem to converge with the study of star employees. The first of these was mentioned previously: high-potential employees. A high potential employee, or Hi-Po employee, is an employee who has the promise or possibility of becoming something more than she is currently (Silzer & Church, 2010).

This could encompass any level of potential; from moving up a step to become a low-level team leader, or even moving into a C-suite role. A high-potential employee is someone who has the ability and promise to be better. Companies are interested in identifying and developing these individuals as part of succession planning and leadership development (Silzer & Church, 2010). While some organizations use past performance as an indicator of an employee being high potential, research notes that current and past performance are not accurate indicators of future performance (Silzer & Church, 2010). This is a first main difference between Stars and high-potential employees. Stars are defined by their superiorly high job performance, while high-potential employees are not. While high-potential employees may have the "potential" to become a star, the two types of employees are not the same. Potential could mean different things in different companies' Hi-Po identification systems. In addition, assessments for Hi-Po's are typically tailored to what the company wants or what position needs filled in the future. This is a fundamental difference between high-potential employees and star employees. There exist several measures meant to identify high-potential employees. These are offered by various assessment organizations and include but are not limited to the following measures: Hogan's High Potential Leadership Reports, Red Bull's Wingfinder Analysis, Korn Ferry Assesment of Leadership Potential, HayGroups Hidden Potential Assesment, Central Test's Professional Profile 2 and CTPI-R, etc. Although I argue that high-potential employees are different from star employees, these existing Hi-Po identification measures could be interesting to examine in relation to star research. Unfortunately, these Hi-Po measures are not available to the public.

High performing employees.

As mentioned previously, sometimes the term star and high-performer are used synonymously. Call, Nyberg, and Thatcher (2015) express that an employee who is a disproportionately high performer may not be considered a star unless they also possess the visibility and social capital aspects of a star. The researchers claim that an employee who fits this description is referred to as a high-performer; therefore, star employees are a subset of high performers. This notion is not completely agreed upon. This disconnect likely results from the lack of an agreed upon definition of a star performer, and well as a lack of understanding surrounding the term high-performers. For example, Trevor, Gerhart, and Boudreau (1997) refer to low, average and high performing employees. In this case a high performer is the top category of three categories of employee's performance levels. However, high performance is a common descriptor when referring to stars (e.g. Groysberg & Lee, 2008).

Experts.

Expertise in another area of study that has overlapping characteristics with star employees. An expert can be defined as "having, involving, or displaying special skill or knowledge derived from training or experience" (Ericcson & Towne, 2013; Merriem-Webster Dictionary, 2009). There are two main approaches to expertise. The first is the Traditional Approach, which defines individual experts by peer nomination or the amount of time they have worked as professionals (Ericcson & Towne, 2013). Expertise has been found to be related to the extensive experience that an expert has obtained (Ericcson & Lehmann, 1996). In fact, in this approach to expertise, experience is the most vital and

necessary aspect of becoming an expert (Ericcson & Towne, 2013). While I believe experience is beneficial in becoming a star, it may not be necessary. Simply put, there is not research to suggest that experience is actually necessary to be considered a star.

The second is the expert-performance approach which defines the individual expert by the reproducible and superior performance within the expert's domain (Ericcson, 2006a, 2006b). This definition has similarities with that of a star employee. Star employees are overarchingly, consistently, high performers, and perform significantly higher than their peers. However, star's have superior *job* performance while experts have superior performance in the domain of their expertise. I am convinced that not all star employees are experts, and not all experts are star employees, though the two may often overlap. Stars could certainly be considered an expert in their domain, however their performance shines through everything they do, not just in activities related to their expertise. On the same hand, not all experts would be considered stars. For example, anyone with a doctoral degree may be considered an expert in their field, but this expertise does not equate to stardom. Some IO psychologists may have star performance while others do not, but they would all be experts in IO psychology. Call, Nyberg, and Thatcher (2015) express the same opinion. However, according to Call, Nyberg, and Thatcher (2015) experts are also a subset of high-performers. This is also not an agreed upon notion, as not all experts can be assumed to be high performers. The problem, again, is the varied use of the term high performer.

In addition to the similar definition, the motivation for studying experts and star employees can be the same. Hypothetically, there are general characteristics of experts that differentiate them from others (Simon & Chase, 1973). This is the same hypothesis that I am applying to star employees. I believe there are some general characteristics of stars that differentiate them from others. A few mediators of expert domain specific performance have been uncovered. Intelligence is a mediator, but only in the beginning stages of learning a skill (Ericcson & Towne, 2013). Perceptual abilities then significantly mediate performance, but this relationship only lasts until a skill has been automated (Ericcson & Towne, 2013). Finally, psychomotor abilities can be helpful after a skill is automated (Ericcson & Towne, 2013). Experts also tend to anticipate, reason, plan, control, and monitor often. The cognitive research in expertise has theorized that experts have "long-term working memory." In other words, they are able to recall a large amount of information from long-term memory at the speed in which working memory occurs. This leads cognitive psychologists to believe that experts utilize better encoding methods, with larger chunks or information (Ericcson & Towne, 2013). Finally, the most important mediator affecting an expert's performance is "deliberate-practice." This is practice in which the person is fully concentrated, gets immediate feedback, makes improvements, and then expands their performance by doing exercises such as performing for longer times without breaks (Ericcson & Towne, 2013). Though experts and stars are not the same, these mediators of expertise could be an important consideration when examining star employees, therefore, should be kept in mind.

The measurement of expert performance has the same issues that the measurement of job performance has. There are objective measures, which measure expertise accurately in something like swimming or running, but do not capture expertise in something like sales dollars. Expertise research attempted to capture performance using discrete independent tasks. These tasks correlated highly with performance but did not fully capture performance. Memory was also tested within the expert's domain, and this task did not capture the knowledge and thinking skills required of an expert.

Why Do Stars Become Stars?

Something this paper has not yet addressed is the question of why stars become stars. Much of the current study focuses on individual differences in performance, visibility, social capital, personality, and behavior as the discriminating or differentiating factor. However, there are other potential theoretical reasons why stars might rise to stardom while others do not.

One possible theory relates to the person-situation debate within personality literature. The premise of this debate is over what accounts for human behavior – the situation a person is in or that person's personality. There is support for both these perspectives. Ultimately, we know that both the situation and personality, as well as the interaction between the two, affect behavioral outcomes. It could be argued that stars were able to become stars not because of their personality but because of the circumstances they were in.

Another relatable theory is often referred to as the Matthew effect (Merton,

1968b). The Matthew effect has been colloquially described as 'the rich get richer and the poor get poorer.' In this way, an employee experiences an advantage of some sort, which begets another advantage and so forth. Perhaps stardom is simply acquired by compiling advantages across time.

Current Research

Call et al. (2015) identified four subcategories of the relationships typically examined in star literature: star formation, organizational effects, colleague effects, and star mobility. Current research on star employees has focused on these four categories, with very little research examining what a star really is or why a person becomes a star. Research has examined how to best manage or produce star performers (Aguinis & Bradley, 2015) as well as what the role of manager of a star performer is (Heslin, 2009). Groysberg, Lee, and Nanda (2008) examined star employee's mobility across organizations. Studies have focused on the best way to retain star employees (Hausknecht, Rodda, & Howard, 2009), as well as the effects of star turnover on firm performance (Kwon & Rupp, 2012). Oldroyd and Morris (2012) studied the effect of social capital information overload on star employees. Groysberg and Lee (2008) examined what effects the quality of a star's colleagues have on that star's performance. Finally, researchers have examined the utility of hiring stars from other companies (Groysberg, Nanda, & Nohria, 2004). Despite this wide field of research regarding stars, there is a very important and fundamental area that researchers have not thoroughly studied: What is a star employee and what makes them different than other good employees?

Very few studies have investigated the behaviors, characteristics, or traits of stars. Kelley and Caplan (1993) found that star performers do not differ from average performers based on cognitive abilities. This could be explained by the sample consisting of all quite intelligent IT professionals, i.e., a floor effect. However, these researchers found that there were nine separate work strategies that seemed to distinguish between star employees and other good employees: taking initiative, networking, self-management, teamwork effectiveness, leadership, followership, perspective, show-and-tell, and organizational savvy. In addition, Kelley and Caplan (1993) found that top performers and middle performers both spoke of these categories, but they differed in how each strategy ranked in importance and how each strategy was described. Star performers rated taking initiative at the highest level, followed by networking, self-management, perspective, leadership, teamwork effectiveness, and followership. Organizational savvy and show-and-tell were of the least importance. This ranking was reversed for the non-star employees. Furthermore, non-star employees would discuss taking initiative at a level that stars considered just part of the job. Stars would consider taking initiative to mean going above and beyond the call of duty. Similar differences were found throughout most strategies.

Bish and Kabanoff (2015) attempted a follow up experiment to provide conceptual clarity about the differences between star performers and non-star performers. These researchers divided the broad construct of performance into two commonly used facets: task performance and contextual performance. Task performance referred to the technical skills and knowledge needed to complete job tasks whereas contextual

performance referred to behaviors that support the organizational context (Borman & Motowidlo, 1993). The researchers sent surveys out to mid-level managers in an Australian public-sector agency. They survey consisted of several qualitative questions with open-ended responses as well as a measure of task and contextual performance. Managers were asked to evaluate whoever they wanted to, and after evaluating them they needed to indicate whether the person they evaluated was a star or not. Bish and Kabanoff (2015) found that although above-average scores on these two dimensions of performance are necessary for an employee to be considered a star, these performance dimensions were not sufficient to distinguish star performers from the non-star performer. The researchers explored additional behaviors and characteristics of stars using manager's descriptions of star employees gathered from the qualitative survey questions. These descriptions were analyzed first by utilizing Braun and Clarke's (2006) process to establish themes. After this, the researchers considered the significance of the themes and focused on the elements not accounted for by the task and contextual performance measures. This process led to the identification of five major categories: team development, self-direction, knowledge and skills, goal achievement, and leadership. These five categories will be used in this study to measure performance. Specifically, levels of performance in these categories will be measured for both stars and non-stars. Based on this study, I would expect stars to be rated significantly higher than non-stars on all five of these categories.

The apparent lack of research regarding what a star is and what makes a person a star gives way to the goals of this study. First, the current study seeks to identify what makes stars different from other good employees. I will quantitatively test the integrative definition of stars established by Call, Nyberg, ad Thatcher (2015). I will examine the differences between stars and non-stars across the major types of performance: global, task, and contextual. I will also examine the performance differences between stars and non-stars across the five domains established by Bish and Kabanoff (2015). On the qualitative side, I will examine managers' perceptions of a star as well as explore and identify additional traits and behaviors that differentiate stars. Together, this information will help us better understand star employees and potential methods of identifying them in the absence of observable criteria.

II. METHOD

Design and Power Analyses

This study has one within subject factor (i.e., star vs non-star ratings). Because there are several possible analyses for this study, several power analyses were run. Power analyses were conducted for a two-tailed medium effect size (i.e., $d = 0.50$), an alpha level of .05, and a power of .80 for both a dependent *t*-test and an independent *t*-test. For a dependent t-test the power analysis indicated 34 participants were required. For the independent t-test the analysis indicated 64 participants were required for both groups. Estimates for sample size needed for the qualitative portion of the study vary, with Morse (1994) suggesting a sample size of 30-50 for grounded theory and Creswell (1998) suggesting 20-30 participants is adequate as a large sample will lead to information saturation. Based on these results this study aimed to collect responses from a minimum of 60 managers. The sample size for each sample collected was smaller than this estimate (i.e., $n = 40$ and $n = 46$); however, upon examining actual effect sizes these sample sizes were found to be satisfactory.

Participants

Participants were recruited using two different methods. The two methods were used to support each other because the target demographic was difficult to collect. The first sample was sourced from Amazon's Mechanical Turk (MTurk) and participants who answered yes to a screening question were allowed to participate. The screening question

was “Do you work in a managerial position supervising other employees?” This question was hidden in several other fake screening questions. After data cleaning, 40 total participants were used. Participants ranged from 26 to 61 years of age while the average age of participants was 38 years (*SD* = 8.6). The majority of participants were male (i.e., 69%) and the majority of participants had a bachelor’s degree (i.e., 64%). Four participants identified themselves as working in manufacturing, six identified as working in retail, one identified as working in transportation and warehousing, four identified as working in information, five identified as working in financial activities, seven identified as working in professional and business services, five identified as working in health care and social assistance, one identified as working in leisure and hospitality, and three identified as working in government. Participants had a range of 2 to 70 subordinates (*M* = 13.1, *SD* = 15.3). Participants had been in their current job anywhere from one year to 36 years (*M* = 7.1, *SD* = 6.5).

The second sample consisted of top executives recruited from several businesses and industries. These participants were recruited from the primary researcher’s social network, and included business owners, CEOs and other c-suite personnel, and top-level managers. Approximately half of the participants were partners at a multi-national, multibillion-dollar corporation in the public services industry. Participation was voluntary. After data cleaning, 46 total participants were used. Participants ranged from 36 to 66 years of age while the average age of participants was 48 years (*SD* = 8.1). The majority of participants were male (i.e., 78%) and the majority of participants had a bachelor’s degree (i.e., 45.7%) while 37% had a master’s degree and 13% had a doctoral

degree. Two participants identified themselves as working in construction, seven participants identified themselves as working in manufacturing, one identified as working in retail, one identified as working in utilities, one identified as working in information, five identified as working in financial activities, 22 identified as working in professional and business services, two identified as working in health care and social assistance, three identified as working in government, and two participants reported working in another industry. Therefore, almost half of this sample (i.e., 47.8%) reported working in professional and business services. Participants had a range of 0 to 320 subordinates (M = 45.8, SD = 59.6). Participants had been in their current job anywhere from four months to 33 years (M = 9.3, SD = 7.8).

These data sources were analyzed separately as they are likely from two different populations: top level executives and managers of any kind. The executive sample had a significantly higher age (M = 48.1, SD = 8.1) than the MTurk sample (M = 37.2, SD = 8.8), t (80) = 5.88, p = <.001, d = 5.88. The participants from the executive data set also had higher education levels than the MTurk sample χ^2 = 16.29; df = 6, p = .012, V = 0.13 (Cramer's V effect size, medium effect). Finally, the executive data set had a significantly higher number of subordinates (M = 45.8, SD = 59.6) than the MTurk sample (M = 13.0, SD = 14.8), t (84) = 3.39, p = .001, d = 3.39. There were no differences in gender between the two samples, χ^2 = 1.08; df = 1, p = .299.

Measures

Measures included in the survey were both quantitative and qualitative. Participants began by reading a description of the study and responding to a manipulation

check. Next participants answered three open-ended questions (i.e., prototypical and specific qualitative questionnaires). Next participants rated employees (both a star and a non-star) on performance, visibility, and social capital measures. Finally, participants provided demographic and job information.

Manipulation check. Participants were instructed to read instructions which included a description of the employees they should be rating during the study. This included a brief description of a star employee, as well as a description of an above-average non-star employee (see appendix for entire text). On the next screen participants were given two questions. The first was "What is meant by *above average non-star employee*?" The three response options included "Any employee that I do not consider a star", "A very good employee that I do not consider a star", and "An average employee". If the participant did not select the correct response (i.e., response option two), they were redirected to the instructions and given a chance to answer the manipulation check again. The second question was simply meant to ensure that the participant had a star employee in mind whom they could rate. The question was "Have you thought of a *star employee* and an *above average non-star employee* in which you can evaluate?". Responses included "Yes, I have both employees in mind" and "No, I don't know any star employees". No participant selected the second option during the study.

Demographic information of participant. The manager was asked to report their own gender, age, and highest level of education completed.

Demographic information of employees. The manager was asked to estimate

the star employee's age, gender, and highest level of education completed. They were also asked to estimate the non-star's age, gender, and highest level of education completed.

Job information. The participant was asked to report the industry in which he works, his own job title, as well as the job title of the star employee and the non-star employee he chose to evaluate. He was also asked to report the approximate number of people who work for him, and the duration that both the star and the non-star employees he chose to evaluate worked for him.

Performance. Performance levels were measured for both the star and the non-star that the participant chose to evaluate. Each was rated on a global level of job performance, an overall level of task performance, an overall level of contextual performance, and five specific dimensions of job performance as identified by Bish and Kabanoff (2015): team-development, self-direction, knowledge and skills, goal achievement, and leadership. These performance measures were on a 7-point scale. The participant compared the individual to all employees within the organization and selected the percentile in which they fell (i.e., Top 50%, Top 40%, Top 30%, Top 20%, Top 10%, Top 5%, Top 1%). Additionally, both the star and non-star were rated on the extent to which they engage in deviant behaviors that harm the organization. This measure is on a five-point scale ranging from always to never. All these measures were simple one-item measures asking the participant to identify the percentile an employee falls into for each type or domain of performance.

Visibility. A measure of visibility was established by creating an initial 8-item

measure and pilot testing the measure on Amazon's Mechanical Turk. Data were analyzed using classical test theory. The final measure is three items ranging from strongly disagree to strongly agree (a = .81). Item one is "When this person does a good job, others in the organization know about it." (r = .63). Item two is "Other employees are familiar with this person's accomplishments." (r = .74). Item three is "This person's performance is visible to others in the organization." (r = .65). The reliability of the measure within the current study was quite similar to the pilot study. From the MTurk data set, star's visibility had a Cronbach's alpha of .85, while non-star's visibility had a Cronbach's alpha of .90. From the executive data set, star's visibility had a Cronbach's alpha of .80, while non-star's visibility had a Cronbach's alpha of .87.

Relevant social capital. A measure of social capital was established by creating an initial 10-item measure and pilot testing the measure on Amazon's Mechanical Turk. Data were analyzed using classical test theory. The final measure is 3 items ranging from strongly disagree to strongly agree (a = .80). Item one is "This person has a strong professional network." (r = .59). Item two is "When a problem arises, this person always knows who to talk to." (r = .66). Item three is "This person has all the right contacts in order to get things done." (r = .71). The reliability of the measure within the current study was quite similar to the pilot study. From the MTurk data set, star's social capital had a Cronbach's alpha of .80, while non-star's social capital had a Cronbach's alpha of .88. From the executive data set, star's social capital had a Cronbach's alpha of .85, while non-star's social capital had a Cronbach's alpha of .60.

Prototypical qualitative questionnaire. The prototypical qualitative

questionnaire consisted of two open-ended questions asking about a general star and non-star. Item one was “Please describe what you believe distinguishes a star’s performance from a non-star’s performance.” This question is meant to capture the initial thoughts a participant has about a star. Item two had two parts. The first was “Can you think of any behaviors that a star employee displays, and a non-star employee does not?” and the second was “Can you think of any characteristics / personality traits that a star employee displays, and a non-star employee does not?” These were treated as one question with one response box so that participants did not have to mentally organize their thoughts into separate behaviors and characteristics, and so that both behaviors and characteristics could be prompted by the question. This question was meant to prompt participants to think beyond their initial thoughts of a star, specifically thinking of behaviors and characteristics.

Specific qualitative questionnaire. The specific qualitative questionnaire asked the participant about the star and non-star they had chosen to evaluate. This question also had two parts. Part one was “Please describe behaviors that the star employee you selected exhibits. Focus on behaviors that you believe makes this person a star” and part two was “Please describe characteristics/personality traits that the star employee you selected exhibits. Focus on characteristics / traits that you believe makes this person a star.” These were treated as one question with one response box so that participants did not have to mentally organize their thoughts into separate behaviors and characteristics, and so that both behaviors and characteristics could be prompted by the question. This question was meant to prompt participants to consider an actual star rather than their

perception of a star in general.

Procedure

All participants. A link to an online questionnaire (which can be seen in Appendix A) was sent to high-level managers selected by the researcher or organization. Participants were informed that the questionnaire was anonymous and were encouraged to email the researcher with any questions they had. First the participants read the informed consent and choose to participate or not to participate. Next the participant read about the purpose of the study and a description of what is intended by star and above average non-star employee. The above-average non-star employee they chose must still be a great employee who is good at their job, but not considered a star. The participant was asked if they were able to think of a star and a non-star employee which they could evaluate. If the participant could not think of a star and a non-star to evaluate they were directed to the prototypical qualitative questionnaire, then to the participant demographics. This completed the survey. Those participants who answered 'yes' they can think of a star and a non-star employee to evaluate completed the entire survey. First, they completed the prototypical qualitative questionnaire. Next, they completed the specific qualitative questionnaire. After this questionnaire, the participant completed the star and non-star performance, visibility, and social capital assessments. This section of the survey was counter-balanced, meaning that approximately half the participants evaluated the star employee first and then the non-star while the other half evaluated the non-star and then the star employee. Following these assessments, participants completed

the job information questions, the participant demographics, and then the star and non-star demographics. This completed the survey. Participants were thanked for completing the study.

MTurk sample. This procedure was the same for the MTurk participants, with few exceptions. MTurk participants found the survey on their own through Amazon's Mechanical Turk. Each participant was paid $1.00 for completing the study. Participants had to answer a short screening questionnaire in which the screener question "Do you work in a managerial position supervising other employees?" was embedded in a series of other irrelevant questions. If participants answered "yes" to the screener questions they were allowed to participate in the study.

III. RESULTS

Data Cleaning

Data were cleaned for quality using listwise deletion. Any participant who did not pass the manipulation check was excluded from participation, and participants whom did not respond as instructed to the qualitative questions were removed from analyses. No participants were removed from the executive sample, and 15 total participants were removed from the MTurk sample. Additionally, for the MTurk data sample a screener question was presented (i.e., "Do you work in a managerial position supervising other employees?") with several other demographic questions to ensure participants qualifications before the survey. MTurk participants whom did not respond affirmatively to the screener question where not allowed to participate in the study.

Data were checked for odd values, errors, and outliers. Outliers were identified as data points with a z-score of greater than 3.0 or less than -3.0. In the MTurk data set, eight outlying values were identified. In the executive data set, seven outlying data points were identified. Analyses were run both with and without these outlying data points. The removal of these outliers affected one of the 11 dependent variables: deviance behaviors. Results will be reported after the removal of these outliers and the effect that the removal of outliers had on deviance behaviors will be discussed throughout.

Data were also assessed for normality and homogeneity. Many variables were found to have a skewed distribution. The variables were skewed in different directions,

making transformed data difficult to interpret. Therefore, the data were also analyzed using the Mann-Whitey nonparametric test for independent samples and the Wilcoxon Signed Ranks test for paired samples. Results of the nonparametric tests were the same as results for the parametric tests, therefore *t*-tests were used despite the nonnormality of the data. All nonparametric results can be found in Appendix B.

Order Effects

All dependent variables were tested for order effects (i.e., whether rating a star or a non-star first made a difference in responses). Independent t-tests were run for all dependent variables between the stars whom were rated first and the stars whom were rated second. This was repeated for non-stars rated first and non-stars rated second. I would expect to find no differences between stars rated at first or second, as well as no differences between non-stars rated at first or second. No significant results were found; therefore, I can conclude that the order in which participants rated stars and non-stars did not affect the participants employee ratings (see Tables 1 and 2 for order effect results).

Quantitative Data Analyses and Results

Two-way mixed-design MANOVA. A two-way mixed-design multivariate analysis of variance (MANOVA) was conducted to examine the effect of employee status (i.e., star or non-star, within-subjects factor) and sample (i.e., MTurk and executives, between-subjects factor) on all dependent variables. Results showed significant within-subjects' effects, Wilks' $\lambda = .097$, Roy's Largest Root = 9.33, F (11, 56) = 47.52, p = <.001, $\eta_p^2 = .903$. Additionally results showed no significant between-subjects effects, Wilks' $\lambda = .843$, Roy's Largest Root = .186, F (11, 56) = 0.95, $p = .506$, $\eta_p^2 = .157$, and

no significant interactions, Wilks' $\lambda = .870$, Roy's Largest Root = .149, F (11, 56) = 0.76, $p = .678$, $\eta_p^2 = .130$ (see Table 3 for summary of results). These results indicate that there are significant differences between ratings for stars and non-stars, but that there are no significant differences between ratings for the MTurk sample or the executive sample. Follow-up within-subjects univariate tests showed significance for every dependent variable assessed. Results can be seen in Table 4. Despite the similarity in results for both the MTurk sample and the executive sample, I will go on to analyze these data sets separately. As I mentioned in the participants section, these samples are likely from two different populations–high level executives and lower level managers. Furthermore, the samples differ on key demographics, including age, education level, and number of subordinates. For these practical reasons I will analyze these data sets separately throughout the rest of this paper.

MTurk sample. The first sample was Amazon's Mechanical Turk. Descriptive statistics can be found in Table 5. The performance measures were analyzed with dependent *t*-tests. The dependent sample *t*-tests were used to evaluate if participants noted a difference in performance when describing both a star and a non-star employee. As this study is explorative in nature, no specific hypotheses were made; however, based on previous literature star employees should be rated higher on all performance variables, including the five categories identified by Bish and Kabanoff (2015). In addition, if Call, Nyberg, and Thatcher's (2015) definition of star's is accurate, stars should be rated as more visible within the organization and as having more relevant social capital. First, stars were rated significantly higher than non-stars on global performance, t (38) = 7.36,

p = <.001, d = 1.18, task performance, t (38) = 6.46, p = <.001, d = 1.03, and contextual performance, t (38) = 7.47, p = <.001, d = 1.20. Additionally, stars were rated significantly higher than non-stars on all five performance facets from Bish and Kabanoff (2015): self-direction, t (39) = 6.58, p = <.001, d = 1.04; leadership, t (39) = 6.36, p = <.001, d = 1.01; goal achievement, t (38) = 5.41, p = <.001, d = 0.87; team development, t (39) = 4.93, p = <.001, d = 0.78; and knowledge and skills, t (38) = 4.57, p = <.001, d = 0.73. Finally, stars were rated significantly higher on internal visibility, t (39) = 4.04, p = <.001, d = 0.64, and were rated significantly higher on relevant social capital, t (39) = 4.01, p = <.001, d = 0.63. After the removal of outliers, non-stars were rated significantly more likely to engage in deviance behaviors, t (35) = 2.96, p = .006, d = 0.49 (see Table 6 for a summary of these results) However, before the removal of outliers, stars and non-stars were not rated differently on deviance behaviors, t (38) = 1.50, p = .141, d = 0.24. Chi-square tests for independence were run to examine differences between stars and non-stars across age, gender, and education as each of these were categorical variables. No significant differences were found between stars and non-stars for age, χ^2 = 1.84; df = 1, p = .175; gender, χ^2 = 1.11; df = 2, p = .574; or education, χ^2 = 5.14; df = 6, p = .527.

Executive sample. The second source of data included high-level executives (c-suite personnel, owners, partners, etc.). The same analyses used with the MTurk sample were used with the executive data. Descriptive statistics can be found in Table 7. Similarly to the MTurk sample, stars were rated significantly higher than non-stars on global performance, t (45) = 14.42, p = <.001, d = 2.13, task performance, t (45) = 9.39, p = <.001, d = 1.38, and contextual performance, t (45) = 13.12, p = <.001, d = 1.93.

Additionally, stars were rated significantly higher than non-stars on all five performance facets from Bish and Kabanoff (2015): leadership, t (45) = 15.07, p = <.001, d = 2.22; self-direction, t (45) = 13.97, p = <.001, d = 2.06; team development, t (44) = 12.77, p = <.001, d = 1.90; goal achievement, t (45) = 12.43, p = <.001, d = 1.83; and knowledge and skills, t (45) = 7.85, p = <.001, d = 1.16. Finally, stars were rated significantly higher on internal visibility, t (43) = 6.11, p = <.001, d = 0.92, and were rated significantly higher on relevant social capital, t (43) = 9.97, p = <.001, d = 1.50. After the removal of outliers, non-stars were rated significantly more likely to engage in deviance behaviors, t (43) = 2.71, p = .010, d = 0.41 (see Table 8 for a summary of these results). However, before the removal of outliers, stars and non-stars were not rated differently on deviance behaviors, t (45) = 1.93, p = .060, d = 0.29. Overall, effect sizes were larger in the executive sample. Chi-square tests for independence were run to examine differences between stars and non-stars across age, gender, and education as each of these were categorical variables. No significant differences were found between stars and non-stars for age, χ^2 = 1.97; df = 1, p = .161 or for gender, χ^2 = 3.19; df = 2, p = .203. However, stars were reported as having a significantly higher education level (i.e., bachelor's and master's degrees) than non-stars (i.e., associate's and bachelor's degrees), χ^2 = 31.70; df = 6, p = <.001, V = 0.26 (Cramer's V large effect size).

Between-subjects comparison effects. It is possible that the results found in these analyses, although not due to the order of ratings, could be due to the comparative nature of the within-subjects design. Therefore, all dependent variables were also tested for between-subjects comparison effects. This analysis was used to determine if

employee ratings may have been inflated or deflated after the first employee was rated. For example, if a participant rated a star first, they may have then rated the non-star by comparing to the ratings they used for the star and deflating those ratings for the non-star. The same could be said for a participant who rated a non-star first. That participant could then rate the star by comparing to the ratings they gave to the non-star and inflating those previous ratings. To test for this potential effect, dependent variables were analyzed with a subset of the data. Instead of comparing all star ratings to all non-star ratings, only the first ratings were used. Therefore, stars which were rated first were compared to the non-stars which were rated first. This would result in an analysis of the between-subject's differences between stars and non-stars. Results were quite consistent with the within-subject's differences between stars and non-stars. All dependent variables showed significance, with the exception of deviance behaviors (see Tables 9 and 10). Interestingly, the effect sizes for these between-subject comparisons were generally larger than the effect sizes found for the within-subjects results (a comparison of these effect sizes can be found in Tables 11 and 12). This might indicate that participants actually rated stars and non-stars as more similar when making a direct comparison between them.

Qualitative Data Analyses

There are both inductive and theoretical qualitative analyses. My approach in this data analysis will be inductive with a basis in grounded theory, meaning that I will not be using any existing theory to guide the coding of this qualitative data. This approach is data driven and therefore may or may not match any previously defined theories or result

patterns from previous literature (Patton, 1990). The present thematic analysis will be at the semantic level, rather than the latent level, therefore all codes will only show patterns of descriptions. The lack of interpretive analysis should yield more consistent coding results (Patton, 1990). However, the results of this inductive analysis may be examined in the context of existing theories *a posteriori*.

Method. The six-step method presented by Braun and Clark (2006) was utilized to conduct this analysis. The first step of this analysis involved the primary researcher becoming familiar with the data and manually creating initial codes. All possible themes/patterns were given initial codes and these codes were listed out and combined into 10 like themes. Following this, the process was again repeated by the primary researcher by creating initial codes on index cards and manually organizing these cards. The theme creation was duplicated four times by four other people. The themes created by these people and the primary researcher were compared and contrasted, giving attention to both internal homogeneity and external heterogeneity (Patton, 1990). A final 14 themes were created, and descriptions were written for each. Two additional codes were created for irrelevant and extraneous information. These themes and descriptions and can be found in Table 13. Following analyses, the primary researcher discovered a missing category which was coded into the extraneous information category. The final themes and codes discovered through the analyses will be thoroughly examined in the discussion.

I used a program called QDA Miner in order to code the data. The 14-theme codebook was written into the program. The primary researcher than coded all responses

to the first qualitative question from the executive data set. Another coder then coded the same responses. These two coders were compared and inter-rater reliability statistics were calculated. Finding these satisfactory, other coders were identified. All coders were trained on how to utilize the program. Each coder was trained by being given a fake response and taught how to highlight sections of text and apply codes. After completely coding this fake response the coders were asked to code various sets of the data. Coders were not able to view the coding provided by other coders. All participants were rated by multiple coders. Four of the six response sets were coded by three different coders. Due to an error, one response set was coded by two different coders and one response set was coded by four different coders. However, the inter-rater agreements for these sets were similar to the inter-rater agreements calculated for the sets coded by three coders, therefore analyses continued as planned. There were five total coders. Table 14 depicts the crossed design of the thematic analysis.

Coding agreement. Individual differences in coding occur no matter how explicit, unambiguous, and precise the coding rules are. There are two main compatibility issues in qualitative coding. The first is the codebook problem. If coders are allowed to create their own codes, the end result can be unanalyzable. Coders can end up with entirely different categories. This problem was mitigated in this analysis by allowing the primary researcher to create a common codebook for all coders. This process was described above. The other problem is the segmentation problem. Commonly, there is an absence of predefined segments – the length and location of text which should be qualified as a single code. The most common solution to this problem is to predefine all

segments to be coded before coding takes place. However, doing so limits the coders ability to interpret each response individually and makes subjective coding much more inflexible. Therefore, segments were not predefined. In this case, a strict agreement between coders for each code and each coded segment is too stringent of a requirement. Therefore, the definition of agreement must be somewhat relaxed.

Inter-coder agreement can be assessed with several different levels of agreement. Code occurrence is agreement calculated on a dichotomous value that indicates the presence of absence of a code, regardless of the number of times the code appears or the codes location. Code frequency is agreement on the number of times a specific code appears in a document regardless of location. It is calculated by comparing the observed frequency per document for each coder. Code importance is agreement on the relative importance of a code. This is calculated by comparing the percentage of words given a specific code across coders. Finally, code overlap is the most stringent form of inter-coder agreement. Coders must agree on the presence, frequency, spread, and location of specific codes. In this analysis I will examine the first level of agreement – code occurrence. In this qualitative analysis the codebook was comprised of personality traits and related behavioral manifestations of those traits. As researchers know from personality research, facet level traits can be related to several broader factors. Traits tend to cross-load. Therefore, there are not succinct and discrete categories for coders to choose from. The code applied to any given text is quite subjective. Combining the nature of coding personality traits with not including predefined segments for coders means that the location and frequency of a code occurring will likely vary from coder to coder. The

important aspect of this analysis is to examine if the code exists in the data.

The simplest measure of agreement is the proportion of agreement out of the total coding's made. However, this measure often yields spuriously high values because it does not take into account chance agreements from guessing. Several adjustment techniques have been suggested, such as: Scott's pi adjustment (Scott, 1955), Cohen's Kappa (Cohen, 1960), Krippendorff's Alpha (Krippendorff, 2004), and the free marginal adjustment which is equivalent to the S Coefficient (Bennett, Alpert, and Goldstein, 1954), the C Coefficient (Jason & Vegelius, 1979), and Brennan and Prediger's (1981) k_n Index (Zwick, 1988). The statistic I used is the free marginal adjustment (or kappa, κ). Kappa provides a standardized index of IRR which can be generalized across studies. This statistic assumes that all categories have an equal probability of being observed and that coder decisions were not influenced by information about the distribution of the codes (Hallgren, 2012). The percent agreement will also be included in the results, although, as stated, this statistic does not account for chance agreements.

Qualitative Data Results

In the qualitative results I will report the inter-rater agreement previously discussed, as well as two different frequency statistics. The first is percent code. This refers to the percentage of all coding associated with the specific code being examined. The second is percent cases. This refers to the percentages of cases (i.e., individual responses by each participant) which contain the specific code being examined.

The distinctions between questions one, two, and three can be differentiated across two factors. The first factor is the presence or absence of a prompt to the

participant to elicit more specific responses. The second factor is prompting the participant to consider an actual example of a star rather than an idea of a star. To be more specific, question one was meant to elicit the very first thoughts someone has when distinguishing a star from a non-star, without prompting the participant in any direction. This should help capture the most initial and raw thoughts about a star. Questions two and three used prompts to elicit more precise beliefs. Both asked the participant to think specifically about behaviors and personality characteristics. Meanwhile questions one and two differed from question three on the second factor. Questions one and two were meant to elicit a more prototypical idea of a star while question three was meant to make the participant think of an actual example of a star. Unfortunately, this second manipulation might or might not have been effective. All three questions were answered *after* the participant knew they needed to choose a star and non-star to rate. Although only question three asked the participant to actually describe an example of a star, the participant could have been thinking of the exemplar star while responding to questions one and two.

MTurk sample.

Question 1. Question one was "Please describe what you believe distinguishes a star's performance from a non-star's performance." Total inter-coder agreement was 82% ($\kappa = .640$). IRRs for each category can be found in Table 15. The category of performance had the highest percentage code at 22.4%. The next highest percent code was initiative at 9.8%. Similarly, the highest percent cases was performance, with 75% of

participants including performance in their responses. The next highest percent cases were tied between initiative and other, with these codes appearing in 47.5% of cases (see Table 16 for all results).

Question 2. Question two had two parts. The first was "Can you think of any behaviors that a star employee displays, and a non-star employee does not?" and the second was "Can you think of any characteristics / personality traits that a star employee displays, and a non-star employee does not?" Total inter-coder agreement was 84% (κ = .681). IRRs for each category can be found in Table 17. The category of performance had the highest percentage code at 12.2%. The next highest percent code was initiative at 9.8%. Similarly, the highest percent cases was performance, with 43.2% of participants including performance in their responses. The next highest percent cases was other at 36.4% and then assertive and initiative at 29.5% (see Table 18 for all results).

Question 3. Question three also had two parts. Part one was "Please describe behaviors that the star employee you selected exhibits. Focus on behaviors that you believe makes this person a star" and part two was "Please describe characteristics/personality traits that the star employee you selected exhibits. Focus on characteristics / traits that you believe makes this person a star." Total inter-coder agreement was 71.8% (κ = .435). IRRs for each category can be found in Table 19. The category of performance had the highest percentage code at 14.8%. The next highest percent codes were initiative at 12.6% and prosocial at 12%. The highest percent cases was initiative, with 79.1% of participants including initiative in their responses. The next highest percent cases was performance at 72.1% (see Table 20 for all results).

Executive sample.

Question 1. Question one was "Please describe what you believe distinguishes a star's performance from a non-star's performance." Total inter-coder agreement was 80% ($\kappa = .600$). IRRs for each category can be found in Table 21. The category of performance had the highest percentage code at 16.9%. The next highest percent code was initiative at 14.6%. Similarly, the highest percent cases were performance and initiative, with 60% of participants including these in their responses. The next highest percent cases were other (41.8%), leadership (34.5%), and perspective (32.7%; see Table 22 for all results).

Question 2. Question two had two parts. The first was "Can you think of any behaviors that a star employee displays, and a non-star employee does not?" and the second was "Can you think of any characteristics / personality traits that a star employee displays, and a non-star employee does not?" Total inter-coder agreement was 85.4% ($\kappa = .708$). IRRs for each category can be found in Table 23. In this data set and question, the category of prosocial had the highest percentage code at 14%. Similarly, the highest percent cases was prosocial, with 48.1% of participants including prosocial in their responses. The next highest percent cases were leadership, passionate, and performance tied at 32.7% (see Table 24 for all results).

Question 3. Question three had two parts. Part one was "Please describe behaviors that the star employee you selected exhibits. Focus on behaviors that you believe makes this person a star" and part two was "Please describe characteristics/personality traits that the star employee you selected exhibits. Focus on

characteristics / traits that you believe makes this person a star." Total inter-coder agreement was 70.6% ($\kappa = .413$). IRRs for each category can be found in Table 25. The category of prosocial had the highest percentage code at 14.7%. The highest percent cases was prosocial and other, with 61.1% of participants including these in their responses. The next highest percent cases was performance and leadership at 50.0% (see Table 26 for all results).

Overall findings. Across all questions and samples, the most used codes were: performance (55.5%), initiative (47.9%), prosocial (36.6%), leadership (32.4%), passionate (28.8%) dependable (26.9%), assertive (25.3%), continuous learner (23.6%), and perspective (22.7%). When looking specifically at question one, which was intended to capture participants initial thoughts of a star, the most mentioned codes were initiative and performance. When looking at question two, which prompted participants to think further about behaviors and traits, continuous learner, leadership, prosocial, assertive, passionate, and problem solver categories had the largest increase. Finally, in question three, which asked participants to think of an actual star, continuous learner, leadership, prosocial, networks, and communication had the largest increases from both question one and two. Overall, this tells me that the first things a manager thinks of a star is a person with high performance and who takes initiative. In addition, managers don't include social skills in their prototype or schemata of a star but acknowledge a star's superior social skills when thinking of an actual star.

IV. DISCUSSION

The first goal of my study was to examine what differentiates stars from other good employees. Results showed that both mid-level managers and high-level executives rate stars significantly higher than non-stars on performance measures. Executives reported a bigger difference between stars and their above average non-star counterparts than the managers on MTurk did. This could be because managers are likely rating employees who are lower down within the company whereas executives are at the top of the company and are likely rating the people directly below themselves. For the MTurk sample, contextual performance had the largest effect size (d = 1.20) followed by global performance, performance in self-direction, task performance, and leadership performance. For the executive sample, leadership performance had the strongest effect size (d = 2.22), followed by global performance, and performance in self-direction. All independent variables measured were found to be significant, although in the future I would not recommend using deviance behaviors as a method of differentiating stars.

The next goal of this study was to test Call, Nyberg, and Thatcher's (2015) definition of stars. If their proposed definition is accurate, stars should be rated higher than non-stars on performance visibility and social capital. As stars were rated higher on all performance measures, visibility, and social capital, I have found support for this definition. However, it should be noted that performance domains showed much stronger effect sizes than visibility or social capital. Clearly performance is the most important

aspect in identifying a star. Visibility and social capital may be the factors that (1) differentiate a star from a high-performer, (2) contribute to a star's superior performance, and (3) enable a star to reach greatness and success.

In addition to finding support for this definition, another goal of this study was to find empirical support for the five categories identified by Bish and Kabanoff (2015): team development, self-direction, knowledge and skills, goal achievement, and leadership. Significant differences were found between stars and non-stars in all five categories the authors identified. More specifically, these results suggest that the two more important of these dimensions are leadership and self-direction. I also had the goal of examining the major types of performance across stars and non-stars. I found that all three types of performance differentiate stars from above average non-stars, with global performance having the largest effect size, followed by contextual performance, and then task performance.

Results of the qualitative analyses were promising. Although qualitative analyses can be messy and difficult, the information gleaned from them can be very informative. First, I will address the categories. The category 'other' was used for extraneous information which did not fit into any other category. Examining the segments which fell into this category, two new categories were found. The first new category was conscientiousness, which consisted of an employee being meticulous, having great attention to detail, being organized, and being a planner. The second new category was adaptability which includes being open-minded and flexible. Additionally, the category of prosocial was split into two categories, one of which absorbed the networks category.

The first is prosocial, which consists of being a team player, helpful, and altruistic. The second is social skills, which consists of being sociable, personable, relating well to others, and being likable.

After these changes, a final comprehensive list of traits which make a star employee different than a non-star was developed. These traits include: adaptable, assertive, communication, conscientiousness, continuous learner, dependable, initiative, innovative, leadership, passionate, performance, perspective, problem solver, prosocial, self-aware, and social skills. Several of these traits are likely related to each other. For example, having high energy (which falls into the assertive category) is likely also related to the passionate category. The categories innovative and problem solver are also likely linked. To provide creative solutions to problems you need both traits. These 16 categories address the final goal of this study: to identify additional traits and behaviors which differentiate stars. These terms paint a picture of what a star "looks like" to others, and, in a very broad sense, answers the question "What is a star?".

Results of the frequency analysis indicated that performance and initiative are important aspects of being classified as a star. Performance and initiative were highest in the first question, which prompted an initial reaction, and in the third question, which prompted the participant to consider a real person. Question two allowed participants to move past performance and initiative and think more deeply about other important traits. The traits with higher percentages of total codes from the MTurk sample included assertiveness, continuous learner, dependable, leadership, passionate, perspective and other. Traits with higher percentages of codes from the executive sample included

assertive, leadership, passionate, and prosocial. However, examining the percentage of codes could mean that participants were more descriptive of this trait or talked about it more. It is possible that more complex traits required more explanation. If I examine the percent cases, or the percentage of participants who included a code in their response, results are slightly different. Still examining the second question, in the MTurk sample, at least a quarter of participants included the categories: assertive, continuous learner, leadership, passionate, and other. In the executive sample, at least a quarter of participants included the categories: assertive, leadership, passionate, perspective, and prosocial.

Results of the qualitative analysis compliment the results found from the quantitative analyses. As previously discussed, Bish and Kabanoff (2015) found the categories of team development, self-direction, knowledge and skills, goal achievement, and leadership. Kelly and Caplan (1993) found that star performers rated taking initiative as the most important, followed by networking, self-management, perspective, leadership, teamwork effectiveness, and followership. One of the most important categories we found was performance. The necessity of performance is further supported by our own quantitative analyses (global, task, and contextual). Additionally, the categories of knowledge and skills and goal achievement established by Bish and Kabanoff fit into our qualitative category of performance, and we found both these categories to be significantly higher for stars.

Moving past performance, our most common category found was initiative. This corresponds with self-direction (from Bish and Kabanoff) which also had one of the

largest effect sizes we found. It also corresponds with taking initiative and self-management, which in Kelley and Caplan's study were ranked first and third by stars respectively. Leadership was another of the most common categories we found. Leadership (from Bish and Kabanoff) also had one of the largest effect sizes we found. In addition, leadership was ranked fifth by stars in Kelley and Caplan's study. We found prosocial and networks to be very important categories. These correspond with networking (ranked second by stars) and teamwork effectiveness (ranked sixth by stars) from Kelley and Caplan's study. Perspective can also be found frequently from the present analyses, was ranked fourth by stars in Kelley and Caplan's study, and was included as an aspect of leadership in Bish and Kabanoff's study. In summary, across this study and the two previous studies, there is strong support for initiative, leadership, perspective, social skills, and performance as attributes of stars.

In this study, we did find one category that stood out as being vital to stars that was not identified in either of the two previous studies. In this study prosocial stood out as one of the most important attributes of stars. Prosocial behaviors included descriptions such as: being helpful to others, empathetic, altruistic, likable, patient, getting along well with others, and relating with others. Prosocial behaviors were described often in the qualitative data and was the third most used code overall with 36.6% cases. In fact, prosocial behaviors were mentioned as often as leadership was. We did not quantitatively exam prosocial behaviors in this study because we had no prior indication of its importance. . However, we did quantitatively examine contextual performance. As previously stated, Borman and Motowidlo (1993) identified four aspects of contextual

performance. First, contextual activities support the organizational, social, and psychological work environment. Second, contextual activities are common to many or all jobs. Third, contextual performance varies with volitional and predispositional variables. And fourth, contextual activities are not role prescribed, and are generally not included in an incumbents list of formal responsibilities. Contextual performance includes behaviors that are supportive of the organization and the people within the organization and these behaviors are done willfully because of the individuals attributes. Prosocial behaviors are very similar to contextual performance, which was found to be one of the most important types of performance in differentiating stars. Prosocial behaviors likely make a person more likable and more visible. It probably helps expand a person's social network. Behaving in a prosocial way influences some of the most important aspects of being a star. This is an attribute not typically thought of as an indicator of stardom, but this study shows that it is a very important star attribute and should be given attention in future research.

Other categories we found which are different from previous findings include adaptable, assertive, communication, conscientiousness, continuous learning, dependable, innovative, passionate, problem solver, prosocial, and self-aware. Of course, these categories are not unrelated. The trait of creative problem solving (innovation and problem solving) and adaptability could very well contribute to superior performance. The trait of prosocial likely relates to contextual performance, networking/social skills, and social capital. Communication, dependable, and assertive categories all likely contribute to outstanding leadership skills.

The results of this study contribute to the minimal amount of literature on star employees. Because this study was comparing stars to other good, high-performing, employees (e.g., global performance on average for stars was rated in the 95th percentile, while non-stars global performance was rated in the 80th percentile) it contributes to understanding what beyond high performance makes stars different. As we continue to learn about why and how stars become stars, we will be able to improve our identification methods, increase star retention, and even predict future stardom. In addition to these contributions, two new measures were created for this study, relevant social capital and visibility. These measures have shown good internal consistency reliability in three different samples and in the future will be subject to validation studies.

Limitations and Future Research

A first limitation to this study lies in the direct comparison nature of the study. Having managers rate both a star and an above average non-star encourages comparisons, and could potential inflate effect sizes that would otherwise not be found. However, our between-subject analyses support our findings, which helps negate this limitation.

A secondary limitation lies in the nature of star employees themselves. Although these employees are spoken of within the literature and within a majority of businesses, a clear definition of what a star is has not been established. Therefore, when a manager is asked to rate a star, it is difficult to know what criteria the manager is using to identify them. In this way managers could have different concepts of what a star employee is. Or there could be different types of stars that have never been differentiated. While discussing the subjective nature of the criterion (star status), it is important to note that

the data used here was all subjective supervisor ratings. When utilizing such a criterion it is important to note two commonly associated issues: same source bias, or common method bias, and halo effect. Same source bias refers to the common variance associated with utilizing the same source when measuring constructs, be that self-report measures or other-ratings (Conway & Lance, 2010). An ideal method should include multiple sources of data to prevent this phenomenon. It is also possible that there is a halo effect taking place in these employee ratings. Essentially this could mean that because a star employee is a great performer in a global evaluation, the supervisor would rate them highly in all individual attributes (Nisbett & Wilson, 1977, p. 250). This halo effect could be a true halo, or could be providing biased, inflated, information about stars (Bartlett, 1983). Including objective data would be extremely beneficial in examining the differences in performance between employees while accounting for any possible halo errors. Even having access to actual performance appraisals would be more beneficial than subjective ratings only.

As previously mentioned, additional validation studies need to be conducted on both newly created measures. Both measures were tested in three separate samples and promising internal consistency reliabilities were found. Another future direction previously indicated in this paper is to perform a content analysis on the various conjectural theoretical descriptions of star employees. These sources can be found online in various formats and written by seemingly qualified individuals, such as business founders and CEOs. These subjective descriptions are an auspicious source for establishing traits and behaviors that may distinguish stars from other employees.

Although it is not a part of this paper, I am in the process of conducting this analysis. Finally, a more parsimonious and complete model of the traits and behaviors we have found should be developed. Each of these categories should be studied to develop a more comprehensive star model.

Ultimately, a study should be conducted directly examining stars and other above average non-stars within an organization. This would, however, require identification of these employees and willingness to participate in a rather large test battery. This battery should include the variables identified in previous literature and variables identified in the current study, as well as some variables that have not been identified as differences between stars and non-stars. Ideally, it would also include a series of content related questions, including the situational circumstances that may have affected the person's career progression. In addition, performance appraisal data would need to be collected.

There are several questions which, in retrospect, I wish I had included in my study. First, I would ask participants how they would define a star. This would have given me direct insight to how managers perceive stars. Second, I would have asked participants how important they believe stars are to the success of the company. Finally, I would ask managers what percentage of their subordinates they would consider stars. This would provide insight to the ongoing performance distribution debate. Performance has been described as both a normal curve and a power law curve (O'Boyle & Aguinis, 2012). If performance is legitimately a normal curve, there should be very few star employees (e.g., < 1%) whereas is performance is distributed as a power law curve there could be more stars in an organization (e.g., 5%). Additionally, if performance is

distributed on a power law curve, there would be very few, if any, unacceptable performers; however, with a normal distribution the unacceptable performers would be as frequent as the star performers. Examining the number of stars reported in an organization would be a natural follow-up to the present study.

In this study, we examined deviance behaviors in a very explorative way. We did not have reason to believe that stars would differ from other high performers in this domain. As previously stated, deviance behaviors were defined as "behaviors that harm the organization". This question is focused on harmful behaviors, so is a measure of counter-productive work behaviors (CWBs). This question did not specify the types of CWBs to focus on (i.e., withdrawal behaviors, aggression, sabotage, theft, etc.). Additionally, deviant behaviors can exist without harming the organization. Unlike many other findings in this study, such as initiative or prosocial behaviors, we see little benefit in exploring deviant behaviors as a differentiator of stars in the future.

A final thought is directed towards the selection of not yet emerged future stars. Exploring this would require a longitudinal study, giving newly hired incumbents a test battery to measure all factors previously associated with star employees an following these employees' career progressions.

Moving forward, I would immediately begin with developing a valid measure of "starness" as well as developing and testing a nomological network that distinguishes "starness" from other similar constructs.

Conclusion

In this study we tested an integrative definition of a star employee for accuracy and found support for the definition. Additionally, we replicated previous findings in the literature and found strong support for these previous findings. We found that there is a real difference between stars and above average employees, presenting strong effect sizes across eight performance measures, as well as visibility and social capital. We also performed qualitative analyses which have brought researchers closer to an understanding of what traits make a star different than an above-average non-star employee. We found that initiative is one of the most important attributes of a star and may not be a construct we generally measure or select for. We also found that prosocial behaviors, authentically caring about other people, is an important attribute of a star. The traits we found should be tested quantitatively against both stars and above-average non-stars to determine if each trait is establishing a distinction between employee types. In the future this information may lead to better selection systems and identification systems for star employees within various industries.

V. REFERENCES

Aguinis, H., & Bradley, K. J. (2015). The secret sauce for organizational success. *Organizational Dynamics,* 44, 161-168.

Aguinis, H., & O'Boyle, E. (2014). Star Performers in Twenty-First Century Organizations. *Personnel Psychology,* 313-350.

Bartlett, C. J. (1983). What's the difference between valid and invalid halo? Forced-choice measurement without forcing a choice. *Journal of applied Psychology*, *68*(2), 218.

Beck, J. W., Beatty, A. S., & Sackett, P. R. (2013). On the distribution of job performance: The role of measurement characteristics in observed departures from normality. *Personnel Psychology*, 67, 531–566.

Bennett, E. M., Alpert, R., & Goldstein, A. C. (1954). Communications through limited-response questioning. *Public Opinion Quarterly*, *18*(3), 303-308.

Bish, A. J., & Kabanoff, B. (2014). Star performers: task and contextual performance are components, but are they enough? *Asia Pacific Journal of Human Resources*, *52*(1), 110-127.

Borman, W. C., & Motowidlo, S. M. (1993). Expanding the criterion domain to include elements of contextual performance. *Personnel Selection in Organizations*; San Francisco: Jossey-Bass, 71.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Brennan, R. L., & Prediger, D. J. (1981). Coefficient kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, *41*(3), 687-699.

Call, M. L., Nyberg, A. J., & Thatcher, S. (2015). Stargazing: An integrative conceptual review, theoretical reconciliation, and extension for star employee research. *Journal of Applied Psychology*, *100*(3), 623.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. *Personnel Selection in Organizations*, 3570, 35-70.

Choo, E. K., Garro, A. C., Ranney, M. L., Meisel, Z. F., & Morrow Guthrie, K. (2015). Qualitative research in emergency care part I: Research principles and common applications. *Academic Emergency Medicine*, *22*(9), 1096-1102.

Conway, J. M., & Lance, C. E. (2010). What reviewers should expect from authors regarding common method bias in organizational research. *Journal of Business and Psychology*, *25*(3), 325-334.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological*

Measurement, *20*, 37-46.

Daum. K. (2015) 11 qualities shared by star employees. *Inc.edu*.

Ericcson, K. A. (2006a). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericcson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.) *The Cambridge Handbook of Expertise and*

Expert Performance, 685-706.

Ericcson, K. A. (2006b). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative task. In K. A. Ericcson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.) *The Cambridge Handbook of Expertise and Expert Performance*, 223-242.

Ericcson K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence on maximal adaptations on task constraints. *Annual Review of Psychology*, *47*, 273-305.

Ericcson, K. A., & Towne, T. (2013). Experts and their superior performance. *The Oxford Handbook of Cognitive Psychology*.

Gagne, F. (2000). Understanding the complex choreography of talent development through DMGT-based analysis. In K. Heller, F. Monks, R. Subotnik, & R. Sternberg (Eds.), International handbook of giftedness and talent (pp. 67–79). New York, NY: Elsevier.

Groysberg, B., & Lee, L. E. (2008). The effect of colleague quality on top performance: The case of security analysts. *Journal of Organizational Behavior*, 29, 1123–1144.

Groysberg, B., Lee, L. E., & Nanda, A. (2008). Can they take it with them? The portability of star knowledge workers' performance. *Management Science*, 54, 1213–1230.

Groysberg, B., Nanda, A., & Nohria, N. (2004). The risky business of hiring stars. *Harvard Business Review*, 82, 92–151.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23.

Harnish, T. (2011). For star employees, hire critical thinkers. *Building Your Team: Hiring and HR, American Express*.

Hausknecht, J. P., Rodda, J., & Howard, M. J. (2009). Targeted employee retention: Performance-based and job-related differences in reported reasons for staying. *Human Resource Management*, *48*(2), 269-288.

Heslin, P. A. (2009). "Potential" in the eye of the beholder: the role of managers who spot rising stars. *Industrial and Organizational Psychology*, *2*(4), 420-424.

Jason, S., & Vegelius, J. (1979). On generalizations of the G index and the phi coefficient to

nominal scales. Multivariate Behavioral Research, 14, 255-269.

Kelley, R., & Caplan, J. (1993). How bell labs creates star performers. *Harvard Business Review*, '*93*(July-August).

Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research, 30*(3), 411-433.

Kwon, K., & Rupp, D. E. (2013). High-performer turnover and firm performance: The moderating role of human capital investment and firm reputation. *Journal of Organizational Behavior*, *34*(1), 129-150.

LaPan, C. (2013). A review of QDA Miner. *Social Science Computer Review, 31*(6), 774-778. DOI: 10.1177/0894439313492711.

Maher. N. Four important traits of star employees. *TJ Peel Technology Recruitment*.

Merton, R. K. (1968a). Social theory and social structure. Simon and Schuster.

Merton, R. K. (1968b). The Matthew effect in science: The reward and communication systems of science are considered. *Science, 159*(3810), 56-63.

Murphy, K. R. (1989). Is the relationship between cognitive ability and job performance stable over time? *Human Performance*, *2*(3), 183-200.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological review*, *84*(3), 231.

Nyberg, A. (2010). Retaining your high performers: Moderators of the performance-job satisfaction-voluntary turnover relationship. *Journal of Applied Psychology*, 95, 440–453.

O'Boyle Jr, E., & Aguinis, H. (2012). The best and the rest: Revisiting the norm of normality of individual performance. *Personnel Psychology*, *65*(1), 79-119.

Oldroyd, J., & Morris, S. (2012). Catching falling stars: A human resource response to social capital's detrimental effect of information overload on star employees. *Academy of Management Review*, *37*(3), 396-418.

Patton, M. Q. (1990). Qualitative evaluation and research methods. SAGE Publications, inc.

Saldaña, J. (2015). The coding manual for qualitative researchers. Sage.

Scott, W. A. (1955). Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly*, 321-325.

Silzer, R., & Church, A. (2010). Identifying and assessing high potential talent; Current organizational practices. *Strategy-driven talent management: A leadership imperative, 2*8, 213-280.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist*, 61, 394-403.

Staffing. T. (2015). The top 15 habits of highly productive employees. *Employee Relations, Hiring, Job Seekers, Management*.

Steinert, L. (2013). 5 traits of a star employee. *Inc.edu*

Tartakovsky, M. (2011). 10 traits of highly productive people. *Psych Central*.

Trevor, C. O., Gerhart, B., & Boudreau, J. W. (1997). Voluntary turnover and job performance: Curvilinearity and the moderating influences of salary growth and promotions. *Journal of Applied Psychology*, *82*(1), 44.

Tubré, T., Arthur Jr, W., & Bennett Jr, W. (2006). General models of job performance: Theory and practice. *Performance Measurement: Current Perspectives and Future Challenges*, 175-203.

Viswesvaran, C. (1993). Modeling job performance: Is there a general factor?. Iowa University, Iowa City.

Zwick, R. (1988). Another look at interrater agreement. *Psychological Bulletin*, *103*(3), 374.

Table 1

MTurk: Independent t-test for Order Effects

	t	*p-value*
Global		
Stars	.788	.436
Non-stars	-.566	.575
Task		
Stars	-.572	.571
Non-stars	-.375	.710
Contextual		
Stars	.423	.675
Non-stars	-.596	.554
Team development		
Stars	.031	.976
Non-stars	.075	.940
Self-direction		
Stars	-.141	.888
Non-stars	-.140	.889
Knowledge and Skills		
Stars	.359	.721
Non-stars	-.473	.639
Goal Achievement		
Stars	.753	.456
Non-stars	-.659	.514
Leadership		
Stars	-.755	.455
Non-stars	-.398	.693
Deviance Behaviors		
Stars	.937	.355
Non-stars	-.564	.576
Visibility		
Stars	.042	.967
Non-stars	-.024	.981
Social Capital		
Stars	.712	.481
Non-stars	.277	.783

Note. *p-values* are two-tailed, df = 38. Values calculated by comparing stars rated first to stars at rated second, as well as non-stars at rated first and non-stars rated second.

Table 2

Executives: Independent t-test for Order Effects

	t	*p-value*
Global		
Stars	.376	.709
Non-stars	-.507	.615
Task		
Stars	.745	.460
Non-stars	-.787	.435
Contextual		
Stars	.533	.596
Non-stars	-.109	.913
Team development		
Stars	1.53	.134
Non-stars	-.326	.746
Self-direction		
Stars	.450	.655
Non-stars	.441	.662
Knowledge and Skills		
Stars	.676	.503
Non-stars	.114	.910
Goal Achievement		
Stars	<.001	1<.001
Non-stars	-.525	.602
Leadership		
Stars	-.187	.852
Non-stars	-.109	.914
Deviance Behaviors		
Stars	.517	.608
Non-stars	.377	.708
Visibility		
Stars	.035	.973
Non-stars	-.425	.673
Social Capital		
Stars	-.314	.755
Non-stars	-.906	.370

Note. *p-values* are two-tailed, *df* = 45. Values calculated by comparing stars rated first to stars at rated second, as well as non-stars at rated first and non-stars rated second.

Table 3

Two-Way Mixed Factors MANOVA Results

	Wilks' λ	*RLR*	*F*	*p-value*	partial η^2
Status	.097	9.33	47.52	<.001	.903
Sample	.843	.186	.945	.506	.157
Status x Sample	.870	.149	.759	.678	.130

Note. *p-values* are two-tailed, RLR is Roy's Largest Root, within-subjects factor is status, between-subjects factor is sample, $N = 86$.

Table 4

Within-Subjects Univariate Test Results

	F	*p-value*	partial η^2
Global	292.62	<.001	.816
Task	129.17	<.001	.662
Contextual	214.31	<.001	.765
Team development	173.56	<.001	.724
Self-direction	194.38	<.001	.747
Knowledge and skills	89.60	<.001	.576
Goal Achievement	201.46	<.001	.753
Leadership	243.14	<.001	.787
Deviance Behaviors	23.32	<.001	.261
Visibility	69.88	<.001	.514
Social Capital	126.33	<.001	.657

Note. *p-values* are two-tailed, RLR is Roy's Largest Root, *N* = 86

Table 5

MTurk: Descriptive Statistics

	M	*SD*
Global		
Stars	6.05	1.28
Non-stars	4.08	1.31
Task		
Stars	5.95	1.19
Non-stars	4.15	1.42
Contextual		
Stars	6.03	1.18
Non-stars	3.88	1.45
Team development		
Stars	5.32	1.54
Non-stars	3.70	1.67
Self-direction		
Stars	5.88	1.67
Non-stars	3.88	1.68
Knowledge and Skills		
Stars	5.79	1.32
Non-stars	4.40	1.46
Goal Achievement		
Stars	5.97	1.04
Non-stars	4.25	1.68
Leadership		
Stars	5.55	1.63
Non-stars	3.63	1.78
Deviance Behaviors		
Stars	4.74	0.60
Non-stars	4.47	0.60
Visibility		
Stars	4.38	0.75
Non-stars	3.80	0.88
Social Capital		
Stars	4.17	0.89
Non-stars	3.53	1.00

Table 6

MTurk: The Differences Between Stars and Non-stars

	t	*df*	*p-value*	*d*
Global	7.36	38	<.001	1.18
Task	6.46	38	<.001	1.03
Contextual	7.47	38	<.001	1.20
Team development	4.93	39	<.001	0.78
Self-direction	6.58	39	<.001	1.04
Knowledge and skills	4.57	38	<.001	0.73
Goal Achievement	5.41	38	<.001	0.87
Leadership	6.36	39	<.001	1.01
Deviance Behaviors	2.96	35	.006	0.49
Visibility	4.04	39	<.001	0.64
Social Capital	4.01	39	<.001	0.63

Note. *p-values* are two-tailed.

Table 7

Executives: Descriptive Statistics

	M	*SD*
Global		
Stars	6.13	0.78
Non-stars	4.00	1.16
Task		
Stars	5.96	0.79
Non-stars	4.37	1.31
Contextual		
Stars	6.11	0.82
Non-stars	3.67	1.33
Team development		
Stars	6.04	0.80
Non-stars	3.54	1.35
Self-direction		
Stars	6.39	0.65
Non-stars	3.87	1.33
Knowledge and Skills		
Stars	5.78	0.87
Non-stars	4.28	1.28
Goal Achievement		
Stars	6.09	0.73
Non-stars	4.04	1.12
Leadership		
Stars	6.20	0.78
Non-stars	3.20	1.34
Deviance Behaviors		
Stars	4.89	0.32
Non-stars	4.67	0.52
Visibility		
Stars	4.56	0.44
Non-stars	3.92	0.61
Social Capital		
Stars	4.57	0.47
Non-stars	3.50	0.60

Table 8

Executives: The Differences Between Stars and Non-stars

	t	*df*	*p-value*	*d*
Global	14.42	45	<.001	2.13
Task	9.39	45	<.001	1.38
Contextual	13.12	45	<.001	1.93
Team development	12.77	44	<.001	1.90
Self-direction	13.97	45	<.001	2.06
Knowledge and skills	7.85	45	<.001	1.16
Goal Achievement	12.43	45	<.001	1.83
Leadership	15.07	45	<.001	2.22
Deviance Behaviors	2.71	43	.010	0.41
Visibility	6.11	43	<.001	0.92
Social Capital	9.97	43	<.001	1.50

Note. *p-values* are two-tailed.

Table 9

MTurk: Independent t-test for Comparison Effects

	t	*df*	*p-value*	*d*
Global	3.68	38	.001	1.19
Task	2.68	38	.011	0.87
Contextual	3.74	38	.001	1.21
Team development	3.16	38	.003	1.03
Self-direction	3.27	38	.002	1.06
Knowledge and Skills	2.27	38	.029	0.74
Goal Achievement	2.98	38	.005	0.97
Leadership	2.56	38	.015	0.83
Deviance Behaviors	1.88	37	.068	0.62
Visibility	2.29	38	.028	0.74
Social Capital	3.62	38	.001	1.17

Note. *p*-values are two-tailed, df = 38.

Table 10

Executives: Independent t-test for Comparison Effects

	t	*df*	*p-value*	*d*
Global	7.70	44	<.001	2.32
Task	4.91	44	<.001	1.48
Contextual	8.03	44	<.001	2.42
Team development	6.36	44	<.001	1.92
Self-direction	9.04	44	<.001	2.73
Knowledge and Skills	5.39	44	<.001	1.63
Goal Achievement	6.76	44	<.001	2.04
Leadership	9.68	44	<.001	2.92
Deviance Behaviors	2.23	44	.031	0.67
Visibility	4.12	44	<.001	1.24
Social Capital	3.91	44	<.001	1.18

Note. *p*-values are two-tailed, $df = 45$.

Table 11

MTurk: Effect Size d Comparisons, Between and Within Subjects

	within-subjects	*between-subjects*
Global	1.18	1.19
Task	1.03	0.87
Contextual	1.20	1.21
Team development	0.78	1.03
Self-direction	1.04	1.06
Knowledge and Skills	0.73	0.74
Goal Achievement	0.87	0.97
Leadership	1.01	0.83
Deviance Behaviors	0.49	0.62
Visibility	0.64	0.74
Social Capital	0.63	1.17

Note. *p*-values are two-tailed, $df = 45$.

Table 12

Executives: Effect Size d Comparisons, Between and Within Subjects

	within-subjects	*between-subjects*
Global	2.13	2.32
Task	1.38	1.48
Contextual	1.93	2.42
Team development	1.90	1.92
Self-direction	2.06	2.73
Knowledge and Skills	1.16	1.63
Goal Achievement	1.83	2.04
Leadership	2.22	2.92
Deviance Behaviors	0.41	0.67
Visibility	0.92	1.24
Social Capital	1.50	1.18

Note. *p*-values are two-tailed, $df = 45$.

Table 13

All Codes and Descriptions Used in the Thematic Analysis

Code	Description
Assertiveness	High-energy, risk-taking, confident, opportunistic, ambitious, active
Communication	Direct and clear communication, breaks things down simply, interactive
Continuous Learner	Intelligent, seeks knowledge, immerses themselves in information, curious
Dependable	Timely, trustworthy, dependable, honest, reliable
Initiative	Doesn't need managed, self-sufficient, self-starter, independent, takes initiative, self-motivated
Innovative	Creative solutions, original ideas
Leadership	Leaderships skills, managers others, motivates others by actions, works well with the team, helps develop the team
Networks	Networks within and outside the company, has valuable contacts, sociable
Other	Description that is relevant but does not fit into another code
Passionate	Loves what they do, enjoys work, involved in work, engaged, positive attitude
Performance	Phenomenal talent or skill, consistently high performing, consistent, goes above and beyond expectations, achieves results, is successful in tasks, exceeds expected outcome
Perspective	Forward/future thinker, big-picture perspective, executive perspective, strategic
Problem Solver	Realistic, analytical, problem solver, change advocate, improves processes
Prosocial	Helps others, empathetic, altruistic, likable, patient, gets along well with others, relates with others
Self-aware	Seeks feedback, self-improves, self-reflective
Unrelated	Irrelevant information (e.g., a non-star does some of these things but not all of them)

Table 14

Crossed Design of the Thematic Analysis

	MTurk Data Set			Executive Data Set		
	Q1	Q2	Q3	Q1	Q2	Q3
Coder 1	X	X	X	X	X	X
Coder 2	X	X				X
Coder 3			X	X	X	
Coder 4	X			X		X
Coder 5		X	X			X

Note. There were three qualitative questions and two data sets, therefore six response sets. All participants were coded within each question from each data set. Each response set was coded by four different coders and each coder (with the exception of the primary researcher) coded three different response sets.

Table 15

MTurk Question 1: Coder Agreement

Codes	% Agreement	Free Marginal
Assertive	87.5	0.750
Communication	92.3	0.846
Continuous Learner	92.0	0.840
Dependable	81.3	0.625
Initiative	69.2	0.385
Innovative	92.6	0.852
Leadership	81.5	0.630
Networks	91.3	0.826
Passionate	88.5	0.769
Performance	74.0	0.480
Perspective	92.6	0.852
Problem Solver	85.7	0.714
Prosocial	90.0	0.800
Self-aware	87.0	0.739
Other	59.0	0.179
Unrelated	76.3	0.526
Total	82.0	0.640

Table 16

MTurk Question 1: Coding Frequencies

Codes	Count	% Codes	Cases	% Cases
Assertive	8	1.8	4	10.0
Communication	19	4.2	6	15.0
Continuous Learner	15	3.3	5	12.5
Dependable	29	6.4	12	30.0
Initiative	44	9.8	19	47.5
Innovative	21	4.7	7	17.5
Leadership	17	3.8	7	17.5
Networks	9	2.0	3	7.5
Passionate	16	3.6	6	15.0
Performance	101	22.4	30	75.0
Perspective	30	6.7	7	17.5
Problem Solver	19	4.2	8	20.0
Prosocial	32	7.1	10	25.0
Self-aware	5	1.1	3	7.5
Other	38	8.4	19	47.5
Unrelated	47	10.4	18	45.0

Table 17

MTurk Question 1: Coder Agreement

Codes	% Agreement	Free Marginal
Assertive	83.9	0.677
Communication	100.0	<.001
Continuous Learner	86.7	0.733
Dependable	88.9	0.778
Initiative	82.8	0.655
Innovative	92.0	0.840
Leadership	78.6	0.571
Networks	97.5	0.950
Passionate	87.1	0.742
Performance	80.0	0.600
Perspective	92.0	0.840
Problem Solver	84.0	0.680
Prosocial	80.4	0.607
Self-aware	90.9	0.818
Other	61.4	0.229
Unrelated	77.9	0.559
Total	84.0	0.681

Table 18

MTurk Question 2: Coding Frequencies

Codes	Count	% Codes	Cases	% Cases
Assertive	33	7.9	13	29.5
Communication	5	1.2	2	4.5
Continuous Learner	30	7.2	12	27.3
Dependable	27	6.5	9	20.5
Initiative	41	9.8	13	29.5
Innovative	14	3.4	5	11.4
Leadership	26	6.2	11	25.0
Networks	2	0.5	2	4.5
Passionate	32	7.7	12	27.3
Performance	51	12.2	19	43.2
Perspective	25	6.0	7	15.9
Problem Solver	17	4.1	8	18.2
Prosocial	19	4.6	10	22.7
Self-aware	3	0.7	2	4.5
Other	32	7.7	16	36.4
Unrelated	60	14.4	18	40.9

Table 19

MTurk Question 3: Coder Agreement

Codes	% Agreement	Free Marginal
Assertive	65.6	0.313
Communication	85.7	0.714
Continuous Learner	67.6	0.353
Dependable	82.4	0.647
Initiative	53.0	0.060
Innovative	95.8	0.917
Leadership	65.7	0.314
Networks	87.5	0.750
Passionate	79.0	0.581
Performance	54.1	0.082
Perspective	84.0	0.680
Problem Solver	85.7	0.714
Prosocial	71.8	0.436
Self-aware	80.0	0.600
Other	48.8	-0.023
Unrelated	92.0	0.840
Total	71.8	0.435

Table 20

MTurk Question 3: Coding Frequencies

Codes	Count	% Codes	Cases	% Cases
Assertive	19	3.8	13	30.2
Communication	27	5.4	8	18.6
Continuous Learner	28	5.6	14	32.6
Dependable	45	9.0	15	34.9
Initiative	63	12.6	34	79.1
Innovative	12	2.4	4	9.3
Leadership	27	5.4	15	34.9
Networks	10	2.0	4	9.3
Passionate	27	5.4	12	27.9
Performance	74	14.8	31	72.1
Perspective	9	1.8	5	11.6
Problem Solver	19	3.8	8	18.6
Prosocial	60	12.0	19	44.2
Self-aware	5	1.0	5	11.6
Other	54	10.8	25	58.1
Unrelated	20	4.0	5	11.6

Table 21

Executives Question 1: Coder Agreement

Codes	% Agreement	Free Marginal
Assertive	82.3	0.645
Communication	100.0	1<.001
Continuous Learner	89.3	0.786
Dependable	76.6	0.531
Initiative	72.7	0.455
Innovative	80.6	0.613
Leadership	79.7	0.595
Networks	89.6	0.792
Passionate	77.1	0.543
Performance	75.5	0.511
Perspective	82.4	0.647
Problem Solver	83.3	0.667
Prosocial	85.7	0.714
Self-aware	78.1	0.563
Other	58.1	0.163
Unrelated	88.1	0.762
Total	80.0	0.600

Table 22

Executives Question 1: Coding Frequencies

Codes	Count	% Codes	Cases	% Cases
Assertive	21	4.0	10	18.2
Communication	9	1.7	4	7.3
Continuous Learner	11	2.1	5	9.1
Dependable	21	4.0	12	21.8
Initiative	77	14.6	33	60.0
Innovative	16	3.0	9	16.4
Leadership	50	9.5	19	34.5
Networks	4	0.8	4	7.3
Passionate	25	4.7	13	23.6
Performance	89	16.9	33	60.0
Perspective	42	8.0	18	32.7
Problem Solver	15	2.8	7	12.7
Prosocial	16	3.0	10	18.2
Self-aware	18	3.4	9	16.4
Other	38	7.2	23	41.8
Unrelated	75	14.2	30	54.5

Table 23

Executives Question 2: Coder Agreement

Codes	% Agreement	Free Marginal
Assertive	89.7	0.795
Communication	84.6	0.692
Continuous Learner	87.2	0.744
Dependable	84.6	0.692
Initiative	84.6	0.692
Innovative	76.9	0.538
Leadership	94.9	0.897
Networks	89.7	0.795
Passionate	79.5	0.590
Performance	79.5	0.590
Perspective	94.9	0.897
Problem Solver	82.1	0.641
Prosocial	82.1	0.641
Self-aware	92.3	0.846
Other	89.7	0.795
Unrelated	74.4	0.487
Total	85.4	0.708

Table 24

Executives Question 2: Coding Frequencies

Codes	Count	% Codes	Cases	% Cases
Assertive	28	7.1	14	26.9
Communication	13	3.3	10	19.2
Continuous Learner	20	5.1	12	23.1
Dependable	14	3.6	9	17.3
Initiative	32	8.1	16	30.8
Innovative	11	2.8	10	19.2
Leadership	32	8.1	17	32.7
Networks	5	1.3	5	9.6
Passionate	33	8.4	17	32.7
Performance	27	6.9	17	32.7
Perspective	22	5.6	13	25.0
Problem Solver	18	4.6	12	23.1
Prosocial	55	14.0	25	48.1
Self-aware	12	3.0	9	17.3
Other	18	4.6	11	21.2
Unrelated	54	13.7	29	55.8

Table 25

Executives Question 3: Coder Agreement

Codes	% Agreement	Free Marginal
Assertive	75.6	0.512
Communication	76.6	0.532
Continuous Learner	78.9	0.577
Dependable	72.3	0.446
Initiative	72.5	0.449
Innovative	79.0	0.580
Leadership	65.7	0.314
Networks	62.5	0.250
Other	67.4	0.347
Passionate	67.4	0.348
Performance	73.0	0.459
Perspective	66.9	0.339
Problem Solver	66.0	0.319
Prosocial	74.1	0.481
Self-aware	72.1	0.442
Unrelated	61.7	0.233
Total	70.6	0.413

Table 26

Executives Question 3: Coding Frequencies

Codes	Count	% Codes	Cases	% Cases
Assertive	72	6.8	20	37.0
Communication	72	6.8	22	40.7
Continuous Learner	64	6.0	20	37.0
Dependable	63	5.9	20	37.0
Initiative	77	7.3	22	40.7
Innovative	22	2.1	10	18.5
Leadership	75	7.1	27	50.0
Networks	24	2.3	15	27.8
Passionate	66	6.2	25	46.3
Performance	86	8.1	27	50.0
Perspective	54	5.1	18	33.3
Problem Solver	26	2.5	12	22.2
Prosocial	156	14.7	33	61.1
Self-aware	29	2.7	11	20.4
Other	136	12.8	33	61.1
Unrelated	37	3.5	20	37.0

Appendix A

Survey Flow and Items

All participants:

Dear Participant,

You are invited to participate in a research study conducted by graduate student, Montana Woolley, and Professor of Psychology, Dr. Gary Burns. For this study, you will be asked to complete a survey to help determine differences in behaviors and traits among employees of different performance levels. There are no known risks for your participation in this research study. The information collected may not benefit you directly. However, the information learned in this study may be helpful to others. Your completed survey will be stored online.

The survey will take approximately 10 minutes to complete.

Taking part in this study is voluntary. By completing this survey you agree to take part in this research study. You do not have to answer any questions that make you uncomfortable. You may choose not to take part at all. If you decide to be in this study you may stop and exit the survey at any time. Please feel free to view and/or print this

page for your records.

The Institutional Review Board (IRB) as approved this study. [Approval number?]

Individuals from the Department of Psychology, the Institutional Review Board (IRB), Office of Research and Sponsored Programs, and other regulatory agencies may inspect these records. In all other respects, however, the data will be held in confidence to the extent permitted by law. Should the data be published, your identity will not be disclosed.

If you have any questions, concerns, or complaints about the research study, please contact: Montana Woolley (email: woolley.9@wright.edu). If you have any questions about your rights as a research subject, you may call the Wright State IRB Office at (937) 775-4462. You can discuss any questions about your rights as a research subject with a member of the IRB or staff. The IRB is an independent committee made up of people from the University community, staff of the institutions, as well as people from the community not connected with these institutions. The IRB has reviewed this research study.

Sincerely,

Montana Woolley

woolley.9@wright.edu

Dr. Gary Burns

gary.burns@wright.edu

Participants who cannot think of a star and non-star to evaluate:

The purpose of this study is to discover the differences between a **star employee** and a **non-star employee**.

A star employee can be referred to with many different names. A few of these are star employees, high-performers, top-performers, difference performers, and game-changers. Regardless of what label you use, these employees are the true rock-stars. They are highly productive and stand out from the other employees as truly amazing.

For this study, I will ask you to choose a **star employee** in your organization. I will also ask you to choose an employee whom you would not consider to be a star in your organization. This **non-star** should still be a great employee and very good at their job. To elaborate, please do not pick a below average or average employee as your non-star. Please pick an employee who is above average, whom you would not consider to be a star employee.

First, you will be asked several open ended, general questions. You will then be asked to answer questions about both employees you have chosen to evaluate. These employees you choose will remain completely anonymous.

Can you think of a star employee and a non-star employee in which you can evaluate?

- Yes I can do that
- **No, I don't know any star employees**

Please answer the following open-ended questions with as many details as possible.

> Please describe what you believe distinguishes a star's performance from a non-star's performance.
>
> Can you think of any behaviors that a star employee displays, and a non-star employee does not?
>
> Can you think of any characteristics / personality traits that a star employee displays, and a non-star employee does not?

Please answer the following questions.

Please select the industry in which you work.

- Goods-producing: Mining

- Goods-producing: Construction
- Goods-producing: Manufacturing
- Services: Utilities
- Services: Wholesale
- Services: Retail
- Services: Transportation and warehousing
- Services: Information
- Services: Financial activities
- Services: Professional and business services
- Services: Health care and social assistance
- Services: Leisure and hospitality
- Services: Government
- Services: Other
- Agriculture, forestry, fishing, or hunting

What is your job title?

Approximately, how many people work for you?

What gender do you identify with?

- Male
- Female

- Prefer not to say
- Other
-

Please enter your age in years.

- Age: _____

What is your highest level of education completed?

- Less than high school
- High school graduate
- Some college
- Associates degree
- Bachelor's degree
- Master's degree
- Doctoral degree

END OF SURVEY

Participants who can think of a star and non-star to evaluate:

The purpose of this study is to discover the differences between a **star employee** and a **non-star employee**.

A star employee can be referred to with many different names. A few of these are star employees, high-performers, top-performers, difference performers, and game-changers. Regardless of what label you use, these employees are the true rock-stars. They are highly productive and stand out from the other employees as truly amazing.

For this study, I will ask you to choose a **star employee** in your organization. I will also ask you to choose an employee whom you would not consider to be a star in your organization. This **non-star** should still be a great employee and very good at their job. To elaborate, please do not pick a below average or average employee as your non-star. Please pick an employee who is above average, whom you would not consider to be a star employee.

First, you will be asked several open ended, general questions. You will then be asked to answer questions about both employees you have chosen to evaluate. These employees you choose will remain completely anonymous.

Can you think of a star employee and a non-star employee in which you can evaluate?

- **Yes I can do that**
- No, I don't know any star employees

Please answer the following open-ended questions with as many details as possible.

Please describe what you believe distinguishes a star's performance from a non-star's performance.

Can you think of any behaviors that a star employee displays and a non-star employee does not?

Can you think of any characteristics / personality traits that a star employee displays and a non-star employee does not?

Please answer the following open-ended question with as many details as possible. Please consider the star employee that you have chosen to evaluate.

Please describe behaviors that the star employee you selected exhibits. Focus on behaviors that you believe makes this person a star.

Please describe characteristics/personality traits that the star employee you selected exhibits. Focus on characteristics/traits that you believe makes this person a star.

For the following questions, please consider the STAR employee you have chosen to evaluate.

Overall, how would you rate this individual's performance at work compared to all employees in the organization?

(7-point scale: Top 50%, Top 40%, Top 30%, Top 20%, Top 10%, Top 5%, Top 1%)

Overall Performance

Task Performance: performance focusing on specific job duties

Contextual Performance: performance on tasks that go above and beyond specific job duties

Rate this individual on the following aspect.

(5-point scale: Always, Most of the time, About half the time, Sometimes, Never)

To what extent does this person engage in deviant behaviors that harm the organization?

Overall, how would you rate this individual's performance at work compared to all employees in the organization on each of these specific aspects?

(7-point scale: Top 50%, Top 40%, Top 30%, Top 20%, Top 10%, Top 5%, Top 1%)

Team Development

Self-direction

Knowledge and skills

Goal achievement

Leadership

Please indicate how much you agree with each statement, thinking about the employee you have chosen to evaluate.

(5-point scale: disagree – agree)

When this person does a good job, others in the organization know about it.

Other employees are familiar with this person's accomplishments.

This person's performance is visible to others in the organization.

Please indicate how much you agree with each statement, thinking about the employee you have chosen to evaluate.

(5-point scale: disagree – agree)

This person has a strong professional network.

When a problem arises, this person always knows who to talk to.

This person has all the right contacts in order to get things done.

For the following questions, please consider the NON-STAR employee you have chosen to evaluate.

Overall, how would you rate this individual's performance at work compared to all employees in the organization?

(7-point scale: Top 50%, Top 40%, Top 30%, Top 20%, Top 10%, Top 5%, Top 1%)

Overall Performance

Task Performance: performance focusing on specific job duties

Contextual Performance: performance on tasks that go above and beyond specific job duties

Rate this individual on the following aspect.

(5-point scale: Always, Most of the time, About half the time, Sometimes, Never)

To what extent does this person engage in deviant behaviors that harm the organization?

Overall, how would you rate this individual's performance at work compared to all employees in the organization on each of these specific aspects?

(7-point scale: Top 50%, Top 40%, Top 30%, Top 20%, Top 10%, Top 5%, Top 1%)

Team Development

Self-direction

Knowledge and skills

Goal achievement

Leadership

Please indicate how much you agree with each statement, thinking about the employee you have chosen to evaluate.

(5-point scale: disagree – agree)

When this person does a good job, others in the organization know about it.

Other employees are familiar with this person’s accomplishments.

This person’s performance is visible to others in the organization.

Please indicate how much you agree with each statement, thinking about the employee

you have chosen to evaluate.

(5-point scale: disagree – agree)

This person has a strong professional network.

When a problem arises, this person always knows who to talk to.

This person has all the right contacts in order to get things done.

Please answer the following questions.

Please select the industry in which you work.

- Goods-producing: Mining
- Goods-producing: Construction
- Goods-producing: Manufacturing
- Services: Utilities
- Services: Wholesale
- Services: Retail
- Services: Transportation and warehousing
- Services: Information
- Services: Financial activities
- Services: Professional and business services
- Services: Health care and social assistance
- Services: Leisure and hospitality
- Services: Government
- Services: Other
- Agriculture, forestry, fishing, or hunting

What is your job title?

Approximately, how many people work for you?

What is the job title of the STAR employee you chose to evaluate?

What is the job title of the NON-STAR employee you chose to evaluate?

Approximately how long has the STAR employee you chose to evaluate worked for you?

- Less than 6 months
- 1-2 years
- 3-5 years
- More than 5 years

Approximately how long has the NON-STAR employee you chose to evaluate worked for you?

- Less than 6 months
- 1-2 years
- 3-5 years
- More than 5 years

For the following demographic questions, please consider the STAR employee you have chosen to evaluate.

Please select his/her age range:

- Under 18
- 18 - 24
- 25 - 34

- 35 - 44
- 45 - 54
- 55 - 64
- 65 - 74
- 75 - 84
- 85 or older
- I don't know

Please select his/her gender:

- Male
- Female
- Other _____
- Prefer not to say
- I don't know

What is his/her highest level of education completed?

- Less than high school
- High school graduate
- Some college
- Associates degree
- Bachelors degree

- Masters degree
- Doctoral degree
- I don't know

For the following demographic questions, please consider the NON-STAR employee you have chosen to evaluate.

Please select his/her age range:

- Under 18
- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 54
- 55 - 64
- 65 - 74
- 75 - 84
- 85 or older
- I don't know

Please select his/her gender:

- Male
- Female

- Other _____
- Prefer not to say
- I don't know

What is his/her highest level of education completed?

- Less than high school
- High school graduate
- Some college
- Associates degree
- Bachelors degree
- Masters degree
- Doctoral degree
- I don't know

Please answer the following demographic questions.

What gender do you identify with?

- Male
- Female
- Prefer not to say
- Other _____

Please enter your age in years.

Age: _____

What is your highest level of education completed?

- Less than high school
- High school graduate
- Some college
- Associates degree
- Bachelor's degree
- Master's degree
- Doctoral degree

END OF SURVEY

Appendix B

Table 27

MTurk: Wilcoxon Signed Ranks Nonparametric Test Results

	Z	*p-value*
Global	-4.73	<.001
Task	-4.69	<.001
Contextual	-4.84	<.001
Team development	-3.92	<.001
Self-direction	-4.75	<.001
Knowledge and Skills	-3.68	<.001
Goal Achievement	-4.17	<.001
Leadership	-4.51	<.001
Deviance Behaviors	-2.68	.007
Visibility	-3.76	<.001
Social Capital	-3.46	.001

Note. p-values are two-tailed, Wilcoxon Signed Ranks test is a nonparametric test used in place of a dependent *t-test*.

Table 28

Executives: Wilcoxon Signed Ranks Nonparametric Test Results

	Z	*p-value*
Global	-5.85	<.001
Task	-5.34	<.001
Contextual	-5.84	<.001
Team development	-5.77	<.001
Self-direction	-5.91	<.001
Knowledge and Skills	-5.13	<.001
Goal Achievement	-5.76	<.001
Leadership	-5.88	<.001
Deviance Behaviors	-2.53	.011
Visibility	-4.77	<.001
Social Capital	-5.56	<.001

Note. p-values are two-tailed, Wilcoxon Signed Ranks test is a nonparametric test used in place of a dependent *t-test*.

Table 29

MTurk: Mann-Whitney U Nonparametric Test Results for Order Effects

	U	*p-value*
Global		
Stars	145.00	.200
Non-stars	179.50	.598
Task		
Stars	181.50	.869
Non-stars	192.00	.866
Contextual		
Stars	181.00	.856
Non-stars	175.00	.519
Team development		
Stars	189.00	.798
Non-stars	195.50	.944
Self-direction		
Stars	176.00	.515
Non-stars	195.50	.945
Knowledge and Skills		
Stars	169.50	.600
Non-stars	191.00	.845
Goal Achievement		
Stars	169.00	.586
Non-stars	184.50	.706
Leadership		
Stars	196.50	.966
Non-stars	182.50	.668
Deviance Behaviors		
Stars	169.50	.649
Non-stars	160.50	.550
Visibility		
Stars	189.00	.796
Non-stars	190.50	.837
Social Capital		
Stars	167.00	.389
Non-stars	197.00	.978

Note. *p-values* are two-tailed Mann-Whitney U test is a nonparametric test used in place of an independent *t-test*.

Table 30

Executives: Mann-Whitney U Nonparametric Test Results for Order Effects

	U	*p-value*
Global		
Stars	241.00	.577
Non-stars	244.50	.649
Task		
Stars	219.00	.275
Non-stars	234.50	.495
Contextual		
Stars	235.50	.482
Non-stars	259.00	.901
Team development		
Stars	183.50	.091
Non-stars	257.50	.875
Self-direction		
Stars	243.50	.608
Non-stars	235.50	.513
Knowledge and Skills		
Stars	227.50	.389
Non-stars	249.00	.726
Goal Achievement		
Stars	260.00	.913
Non-stars	245.00	.658
Leadership		
Stars	262.50	.962
Non-stars	259.00	.901
Deviance Behaviors		
Stars	240.50	.602
Non-stars	246.00	.844
Visibility		
Stars	246.00	.867
Non-stars	223.00	.484
Social Capital		
Stars	216.50	.532
Non-stars	225.50	.382

Note. p-values are two-tailed Mann-Whitney U test is a nonparametric test used in place of an independent *t-test*.

Table 31

MTurk: Mann-Whitney U Nonparametric Test Results for Comparison Effects

	U	*p-value*
Global	62.50	<.001
Task	99.00	.006
Contextual	65.50	<.001
Team development	82.00	.001
Self-direction	72.50	<.001
Knowledge and Skills	104.00	.009
Goal Achievement	82.00	.001
Leadership	99.50	.007
Deviance Behaviors	138.00	.074
Visibility	104.00	.009
Social Capital	80.00	.001

Note. *p-values* are two-tailed, Mann-Whitney U test is a nonparametric test used in place of an independent *t-test*.

Table 32

Executives: Mann-Whitney U Nonparametric Test Results for Comparison Effects

	U	*p-value*
Global	32.50	<.001
Task	81.50	<.001
Contextual	25.50	<.001
Team development	47.50	<.001
Self-direction	18.00	<.001
Knowledge and Skills	71.50	<.001
Goal Achievement	46.00	<.001
Leadership	17.00	<.001
Deviance Behaviors	193.50	.030
Visibility	104.00	<.001
Social Capital	110.50	<.001

Note. *p-values* are two-tailed, Mann-Whitney U test is a nonparametric test used in place of an independent *t-test*.